AOC 10-7

POPQAE 50 (1) 1–142 (1986)
ISSN 0033–362X

Public Opinion Quarterly

Symposium on Polls: Is There a Crisis of Confidence?

Rating the Polls: The Views of Media Elites and the General Public — *Andrew Kohut*

Evaluating Polls with Poll Data — *Burns W. Roper*

Trends in Attitudes Toward and Participation in Survey Research — *Stephen Schleifer*

Surveys on Surveys: Limitations and Potentialities — *John Goyder*

Gender Differences in Policy Preferences: A Summary of Trends from the 1960s to the 1980s — *Robert Y. Shapiro and Harpreet Mahajan*

Preelection Polls in Israel: Structural Constraints on Accuracy — *Jacob Shamir*

Mass Communication and Political Socialization: Specifying the Effects — *Gina M. Garramone and Charles K. Atkin*

Not Forbidding Isn't Allowing: The Cognitive Basis of the Forbid-Allow Asymmetry — *Hans-J. Hippler and Norbert Schwarz*

Interviewers' Voices and Refusal Rates in Telephone Surveys — *Lois Oksenberg, Lerita Coleman, and Charles F. Cannell*

A Sample Size Formula for Multiple Regression Studies — *Sande Milton*

The Polls: ABM and Star Wars — *Thomas W. Graham and Bernard M. Kramer*

IN MEMORIAM • NEWS AND NOTES • BOOK REVIEWS

SPRING 1986

Public Opinion Quarterly

JOURNAL OF THE AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

AAPOR ADVISORY COMMITTEE

Leo Bogart — Burns W. Roper
Norman M. Bradburn — Bernard Roshco
Andrew Kohut and Eleanor Singer, councillors-at-large

Editor: Eleanor Singer — *Managing Editor:* Ina C. Jaeger

Department Editors:

Philip A. Harding, *News and Notes*
Marc B. Glassman, *Book Reviews*
Phyllis M. Endreny, *Book Notes*
Tom W. Smith, *The Polls*

EDITORIAL BOARD

James R. Beniger
Joan S. Black
Norman M. Bradburn
Charles F. Cannell
Muriel G. Cantor
Albert H. Cantril
Donald Dillman
Martin R. Frankel
Herbert J. Gans
Albert E. Gollin
Robert Hornik
Mary R. Jackman
F. Thomas Juster
Ronald C. Kessler
Mark R. Levy
Philip Meyer
J. Ronald Milavsky
Warren J. Mitofsky
Elisabeth Noelle-Neumann
Stanley Presser
Donald F. Roberts
Morris Rosenberg
Charles R. Wright
Daniel Yankelovich

Volume #50, 4 issues (1986)

© 1986 by the American Association for Public Opinion Research.

Published quarterly: Spring, Summer, Autumn, and Winter, by The University of Chicago Press, Chicago. Second-class postage paid at Chicago, IL. and at additional mailing offices. Subscription rates, U.S.A.: institutions 1 year $38.00; individuals 1 year $19.00. AAPOR member subscription rate, U.S.A.: individuals 1 year $17.00. Student-subscription rate (with copy of validated ID), U.S.A.: individuals 1 year $17.00. Other countries add $3.00 for each year's subscription to cover postage.

POSTMASTER: Please send all change of address notices to Public Opinion Quarterly, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Mail orders and direct subscription information to: Journals Fulfillment Department, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Single copy and back volume information furnished upon written request. Claims for missing issues can be honored only up to three months from date of issue. Indexed by Public Affairs Information Services, Social Sciences Index, United States Political Science Documents, Combined Retrospective Index Sets (CRIS), and Political Science Abstracts (PSA). Available from University Microfilms International (UMI).

Editorial Office: Journalism Building, Columbia University
116th Street and Broadway, New York, New York 10027
Business, Subscriptions, Back Issues: The University of Chicago Press,
5801 South Ellis, Chicago, IL 60637
AAPOR: Post Office Box 17, Princeton, New Jersey 08540

Subscribers and members of AAPOR are urged to notify the Subscription Office as soon as possible of changes of address. In doing so, give the old as well as the new address.

PUBLISHED BY THE UNIVERSITY OF CHICAGO PRESS

Public Opinion Quarterly

VOLUME 50 SPRING 1986 NUMBER 1

Is There a Crisis of Confidence?

Rating the Polls: The Views of Media Elites and the General Public *Andrew Kohut* 1

Evaluating Polls with Poll Data *Burns W. Roper* 10

Trends in Attitudes Toward and Participation in Survey Research *Stephen Schleifer* 17

Surveys on Surveys: Limitations and Potentialities *John Goyder* 27

Gender Differences in Policy Preferences: A Summary of Trends from the 1960s to the 1980s *Robert Y. Shapiro and Harpreet Mahajan* 42

Preelection Polls in Israel: Structural Constraints on Accuracy *Jacob Shamir* 62

Mass Communication and Political Socialization: Specifying the Effects *Gina M. Garramone and Charles K. Atkin* 76

Not Forbidding Isn't Allowing: The Cognitive Basis of the Forbid-Allow Asymmetry *Hans-J. Hippler and Norbert Schwarz* 87

Interviewers' Voices and Refusal Rates in Telephone Surveys *Lois Oksenberg, Lerita Coleman and Charles F. Cannell* 97

A Sample Size Formula for Multiple Regression Studies *Sande Milton* 112

In Memoriam: Herbert H. Hyman, 1918–1985 *Paul B. Sheatsley* 119

News and Notes *Luis C. Torres-Bohl, Guest Editor* 121

The Polls: ABM and Star Wars: Attitudes Toward Nuclear Defense, 1945–1985
Thomas W. Graham and Bernard M. Kramer 125

Book Reviews
Marc B. Glassman, Editor 135

Herbert McClosky and John Zaller, *The American Ethos: Public Attitudes Toward Capitalism and Democracy*
Lee Sigelman 135

Rita J. Simon, *Public Opinion and the Immigrant, Print Media Coverage, 1880–1980* *Ronald Humphrey* 137

John P. Murray and Gavriel Solomon, Eds., *The Future of Children's Television* *Steven T. McDermott* 138

Book Notes *Phyllis M. Endreny, Editor* 140

NOTICE TO CONTRIBUTORS

Three copies of each article should be submitted. Authors are urged not to submit original typescripts. Copies of manuscripts which are reviewed become the property of the journal, and unused copies will be returned only if the author provides a stamped, self-addressed envelope. Bibliographic-references, footnotes, and tables should be typed to conform to the printed style of this journal.

All submissions that make use of survey data not archived at the Inter-University Consortium for Political and Social Research or the Roper Center must contain the following information: (1) A definition of the population from which the sample was drawn and a statement of the selection procedure sufficient to permit at least approximate replication. (2) The response rate and details of its calculation. (If response rate is not appropriate, the number of refusals.) (3) Dates the survey was conducted. (4) Exact wording of all questions used in the analysis.

Upon acceptance of an article by the journal, the author(s) will be asked to transfer copyright in the article to the publisher. This transfer will insure the widest possible dissemination of information under the U.S. Copyright Law.

Public Opinion Quarterly is hospitable to all points of view, provided that the material presented will help to illuminate problems of communication and public opinion. The American Association for Public Opinion Research assumes responsibility only for material signed by an officer of the Association in his or her official capacity. Statements of fact and opinion are made on the responsibility of the authors alone and do not represent the views of the editors.

Public Opinion Quarterly has been registered with the Copyright Clearance Center, Inc. Consent is given for copying of articles for personal or internal use, or for the personal or internal use of specific clients. This consent is given on the condition that the copier pay through the Center the per-copy fee stated in the code on the first page of each article for copying beyond that permitted by the U.S. Copyright Law. If no code appears on an article, the author has not given broad consent to copy and permission to copy must be obtained directly from the author. All articles published prior to 1978 may be copied for a per-copy fee of $1.00, also payable through the Center. The consent does not extend to other kinds of copying, such as for general distribution, resale, advertising and promotional purposes, or for creating new collective works.

Is There a Crisis of Confidence?

For the May 17, 1985 plenary session of the AAPOR Annual Conference, Andrew Kohut and Burns W. Roper were invited by Conference Chair Eleanor Singer to address a question often encountered during the preceding year: Is there a crisis of confidence? Both generously responded to the invitation by fielding special questions bearing on that topic in the Spring of 1985. Their reports, first presented at the 1985 Conference, appear below. Two other papers given at the 1985 Conference bear very directly on the "crisis of confidence" theme, and they are also included in this special section of the *Quarterly*. One, by Stephen Schleifer, shows very clear convergences with Roper's data and conclusions. The other, by John Goyder, raises the question, also acknowledged by the other writers, of how it is possible to use survey research to measure people's attitudes to surveys.

Rating the Polls: The Views of Media Elites and the General Public

ANDREW KOHUT

The credibility of opinion polls was very much at issue during the 1984 campaign. Throughout the year, pollsters and news organizations that conduct polls were subject to a steady stream of criticism about their findings and methods. In a major journalistic treatment of polls in 1984, Barry Siegel of the *Los Angeles Times* wrote: "These examples illustrate something most polling experts know but don't always acknowledge. It is remarkably difficult to measure public opinion accurately" (Siegel, 1984).

As late as five days before one of the most lopsided elections in history, the *Washington Post* wrote: "In these last days before the

---

Abstract To measure the credibility of the polls in 1984, Gallup conducted two surveys, one of the press and one of the general public. Results of the press survey indicated a favorable reaction to the accuracy of the preelection polling in 1984, although there was a division of opinion about whether the polls were better at explaining the vote than predicting it, and whether the polls enhanced or interfered with the electoral process. The public's perception of the polls was a generally positive one both as to their accuracy and their overall effect.

Andrew Kohut is President of The Gallup Organization.

Public Opinion Quarterly Vol. 50:1–9 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-1/$2.50

election, Democratic presidential nominee Walter F. Mondale has hoisted a new banner: 'The polls are dead wrong!' In one sense, he is dead right. Throughout this political year, the variation in poll results has been so pronounced as to call into question the scientific precision of the popular art of opinion-taking" (Haywood, 1984).

While there was variability in the polls, the root cause for lack of confidence in the polls in 1984 lay in their failure to detect Reagan's big victory in 1980. For more than two decades most of the national polls had produced final figures that were either right on target or well within sampling error; 1980 came as a big shock to the public as well as to the pollsters themselves. The polls had failed to live up to expectations, and as a consequence, in 1984 the level of expectations was much lower; many expected the polls to perform badly again.

To some extent this expectation may have been intensified by wishful thinking on the part of Mondale supporters. There was a great deal of discontent with the intrusiveness of polls in the electoral process. As the *Washington Post* quote above attests, in the final months Mondale was certainly running against his deficit in the polls as well as against Reagan.

How well the preelection polls do is significant not only for the small yet growing number of organizations that do public preelection polling but also for the survey research profession at large. The national preelection polls are the most visible expression of the validity of the opinion survey method. If the polls had been as far off the mark in 1984 as they were in 1980, the impact on polling organizations and, more generally, on the survey research profession would have been very severe.

As it turned out, the performance of the polls in 1984 was variable. Some came quite close to the actual division of the vote while others greatly exaggerated the Reagan victory and a few underestimated it. Nonetheless, postelection news coverage gave the national polls generally good reviews, putting less emphasis on their variability and more on the fact that unlike the polls of 1980, they mostly all showed a big Reagan victory in their final result (Claffey, 1984; Rosenbaum, 1984).

Given the extent to which the credibility of the polls was at issue over the course of the 1984 election, a Gallup effort was mounted on behalf of the 1985 AAPOR annual conference to measure reaction to the performance of the polls subsequent to the national election. Two Gallup surveys were conducted—one, a survey of the press (the principal client of national polling organizations); the other, a small survey of the general public, the ultimate audience for published opinion polls. The overriding objective of these surveys was a postelection assessment of image. More specifically, the studies sought to determine what

Table 1. Meeting Expectations

Many news organizations made use of preelection polls in their 1984 election coverage. For each of the pairs of opposites listed below, please indicate your opinion of preelection polls in 1984. (If you feel the item [at the top] completely describes preelection polls, so indicate by circling "1." If you feel the item [below] completely describes preelection polls, so indicate by circling "4." You may, of course, indicate you opinion by circling any number between 1 and 4.)

	All Respondents %		*Executive Managing Editor* %	*Other Editors, Reporters* %
Did a better job than I expected				
1	10	68	5	$\boxed{17}$
2	58		61	53
3	20	26	23	16
4	6		3	8
Did not do as good a job as I expected				
No answer	6		8	6
	100		100	100
N responses	(246)		(65)	(113)

sort of impression was left on the public and how the polls were judged by members of the news media, who presumably have more informed and experienced viewpoints of preelection surveys.

The Gallup survey of the press consisted of 246 questionnaires administered to members of the news media.[1] The poll was conducted by mail in January 1985. The survey, which was part of an ongoing Gallup program, covered a sampling of senior level editors and reporters as well as publishers and management executives of national and major local market news organizations. Of the 246 questionnaires, 137 were completed by print journalists and 104 by broadcast journalists. The majority, 190, represented editors, broadcast producers, and reporters; 56 were publishers or station managers.

Overall, the results of the press survey indicated a favorable reaction to preelection polling in 1984. First, the performance of the published polls exceeded the expectations of news media respondents (Table 1). The style of question was semantic differential, and most—68 percent—chose the categories nearest the better-than-expected alternative. There was little variation by characteristics or attitudes of respon-

---

[1] The initial mailing was made to 415 potential respondents and followed up with one mail prompt to encourage participation.

Table 2. The Accuracy of Polls

	All Respondents %		Type of Media Print %	Broadcast %
Accurate				
1	33	83	29	$\boxed{39}$
2	50		54	45
3	13	14	15	10
4	1		1	2
Inaccurate				
No answer	3		1	4
	100		100	100
N responses	(246)		(137)	(104)

Table 3. Distinctions Among National Polls

%	All Respondents		Type of Media Print %	Broadcast %
All national polling organizations did about the same quality job				
1	11	30	7	15
2	19		19	19
3	30	68	29	32
4	38		$\boxed{43}$	32
Some national polling organizations did a better job				
No answer	2		2	2
	100		100	100
N responses	(246)		(137)	(104)

dent, except that senior reporters and middle-level editors were slightly more emphatic in this opinion than their superiors.

Most saw the preëlection polls as accurate (Table 2). While this was the overall tenor of opinion, it was somewhat more strongly expressed by media respondents who cover *politics*, and by broadcast-affiliated respondents, especially those who work in TV. The perceived accuracy of polls notwithstanding, most respondents felt that some national organizations did a better job than others. Print journalists expressed this view somewhat more strongly than broadcast journalists (Table 3).

Table 4. Forecasting vs. Explaining

	All Respondents %		Audience National %	Audience Local %
Better at explaining why people vote than forecasting elections				
1	14	43	10	15
2	29		20	31
3	25	54	21	26
4	29		41	27
Better at forecasting elections than at explaining why people vote				
No answer	3		8	1
	100		100	100
N responses	(246)		(49)	(193)

Table 5. Effect on Process

	All Respondents %		Voted for: Reagan %	Voted for: Mondale %
Interfere with election process				
1	18	53	11	26
2	35		31	36
3	35	45	41	29
4	10		15	6
Enhance the election process				
No answer	2		2	3
N responses	(246)		(120)	(102)

Surprisingly, there was a close division of opinions as to whether the polls were better at explaining the vote or better at forecasting it, with the balance of opinion on the side of the latter (Table 4). Those employed by media with a national audience were more inclined to see polls as doing a better job of forecasting.

However, the sharpest and most troubling division of opinion was over whether polls *enhanced* or *interfered* with the electoral process (Table 5). Just about equal numbers of respondents took either side of this issue. On this measure there were some very significant variations

Table 6. Exposure to Public Opinion Polls

Do you follow the results of any public opinion poll regularly in any newspaper or magazine?

	1944 %	*1985* %
Regularly	9	$\boxed{25}$
Occasionally	19	16
No	28	59
Never heard of	44	NA
	100	100
N responses	(2000)	(1023)

in opinion. The following were more apt to take the view that polls *enhance* the process—*publishers, station managers,* broadcasters, and those who work in local markets. Feeling that polls interfere with the process was greatest among respondents from the *West,* those who write for *print,* national media, and senior level executives. However, all of these relationships aside, the single greatest correlate of opinion was partisanship—whose who voted for Mondale were far more inclined to see polls as interfering with the process than were those who voted for Reagan.

The survey of the general public was a limited effort. In March 1985, Gallup administered to a national telephone sample of 1120 adult men and women a series of questions about opinion polls which were first addressed to a national sample in 1944. The historic survey was conducted by Hadley Cantril's Princeton office of Public Opinion Research and initially reported by Eric Goldman in *Public Opinion Quarterly* (Goldman, 1944-45).

The Cantril survey provided a rich perspective on current opinion. In 1944 regularly conducted public national opinion polls were barely a decade old, and it was four years before their major preelection blunder. Nonetheless, many questions were quite appropriate to the issue of credibility in 1985, and here was an opportunity to see to what extent public opinion had changed over virtually the lifetime of the polling profession.

Considering opinions in 1944, a slim majority (56 percent) had even heard of public opinion polls and only 9 percent of the population reported that they followed them regularly. Forty years later, the percentage of the population that follows public opinion polls is much greater (Table 6).

Regarding the accuracy of polls, in 1944 a majority of those aware of polls felt that they were pretty nearly right most of the time in making

Table 7. The Accuracy of Preelection Polls

Some polling organizations make frequent predictions of election results. What is your general impression of how well they do—do you think they are pretty nearly right most of the time, or do you think their record is not very good?

	1944[a] %	*1985* %
Right most of the time	57	$\boxed{68}$
Not very good	21	20
Don't know	22	12
	100	100
N responses	(1120)	(1023)

[a] Based on those aware of polls.

Table 8. The Accuracy of Polls on Issues

Do you think poll returns on matters not dealing with elections, but with public opinion toward such things as labor problems or international affairs, are usually pretty nearly right or not right at all?

	1944[a] %	*1985* %
Pretty nearly right	52	55
Not right at all	12	$\boxed{21}$
Don't know	36	24
	100	100
N responses	(1120)	(1023)

[a] Based on those aware of polls.

election predictions. Opinion today is not very different. However, a larger percentage of those questioned in 1985 felt that polls are pretty nearly right most of the time (Table 7).

In 1944, compared to election predictions, it was less clear to the public whether the polls were accurate in dealing with issues. Note that over one in three had no opinion, but of those expressing a view, a *majority* felt the polls were pretty nearly right on such things as labor problems and international affairs (Table 8). Forty years later, the public is also less certain about the reliability of polls on issues, relative to election predictions, and among those expressing an opinion, the proportion saying "Not right at all" has increased significantly.

Finally, in 1944, the overwhelming majority of respondents judged polls as good for the country—but one in five had no opinion (Table 9). Forty years later, an equally large proportion felt polls are good things, but the percentage reporting they are bad things had increased somewhat.

Table 9. Overall Opinion Toward Polls

In general, would you say that polls of the opinions of the public are a good thing or a bad thing in our country?

	1944[a] %	*1985* %
Good thing	73	76
Bad thing	6	$\boxed{12}$
Don't know	21	12
	100	100
N responses	(1120)	(1023)

[a] Based on those aware of polls.

The results of these two surveys suggest that despite frequent criticism of the polls during the 1984 election, elite and public opinion of the polls was generally positive. There was certainly no evidence of a crisis in confidence. Criticism of the polls during the election either failed to reach large proportions of the public and press or was dissipated by the fact that the outcome of the election bore general resemblance to what most polls were showing. One can only speculate that had Reagan's margin of victory been significantly different from what most polls showed in their final days, the results of these surveys might have been very different.

From a historical viewpoint, opinion of the polls today among the public is not very different from what it was more than 40 years ago, when public polls were in their infancy. Unfortunately, there are no data points in between to demonstrate whether or how public opinion about the polls has varied over the past four decades.

The press survey did point to significant concerns among media respondents about the intrusiveness of polls in election campaigns, while the public survey did not cover this area. Supporters of Mondale had fewer positive things to say about the polls than did Reagan supporters.[2]

There are two other important qualifications to these positive findings for polls. Establishing the credibility of polls with a survey begs the question of whether nonrespondents to either survey had less favorable attitudes than those who participated. However, Gallup experience indicates that nonresponse to surveys is usually based on practical matters such as convenience, timing, interviewer approach, and so on. Opinions about polls, privacy, or other substantive criticism of surveys have tended to be relatively minor factors. It should also be noted that if these surveys were, in fact, subject to some bias because of selective participation, the bias would have had to be enormous to change the overall conclusions.

---

[2] For example, compared to Reagan critics, those who approved President Reagan were more likely to feel that polls are pretty nearly right most of the time (by a margin of 72 percent to 64 percent). A similar pattern was evident in each evaluative question.

A more troubling qualification to the press poll is the fact that the margins of favorability are wider than they are deep. As the tables above indicate, few respondents selected the most positive category in each of the interrogations about opinion polls. It is probable that the favorable attitudes expressed by press respondents could be easily changed, given the small proportion who expressed a high degree of confidence. It is also likely that had the poll been conducted prior to the election, results might have been quite different, given the softness of opinion.

To conclude, the surveys should provide a degree of comfort about the public perception of the reliability of opinion polling. On the heels of the 1984 election a positive impression existed; how well it holds up will depend in part on how well polling organizations perform in subsequent elections. The comments of the founders of the British and American Gallup Polls are quite relevant to this point:

Henry Durant:

> My advice to all would-be election forecasters is to think again. It is the most stupid job you can ever take up, no matter how hard you try to find a worse one. If you get the election right, everyone takes it for granted. If you get it wrong, you are standing alone and utterly ashamed, and there is nothing you can do about it.

George H. Gallup:

> It would be difficult to find any other profession whose leaders are held publicly accountable for their actions in quite the same way. However, it is the awful penalties for being "wrong" that have spurred the development of research procedures. As a result, in no other field of the social sciences has the prediction of human behavior reached the same high level of accuracy (Gallup, 1972).

References

Claffey, Charles E.
1984 "Pollsters did better, Gallup did best." Boston Globe, Nov. 11.
Gallup, George
1972 The Sophisticated Poll Watcher's Guide, 1972. Princeton, NJ: Opinion Press.
Goldman, Eric F.
1944–45 "Poll on the polls." Public Opinion Quarterly 8:461–67.

Haywood, Richard
1984 "Conflicting campaign polls point to one certainty: some are wrong." Washington Post, Oct. 31.
Rosenbaum, David E.
1984 "A good election for poll takers." The New York Times, Nov. 8.
Siegel, Barry
1984 "Polls: more art than science." Los Angeles Times, Aug. 27.

Evaluating Polls with Poll Data

BURNS W. ROPER

In an attempt to learn how the American public feels about opinion polls, The Roper Organization conducted a national poll in late March of 1985.[1] A total of 10 questions was asked. First people were asked if they had been interviewed before "this" interview, either by telephone or in person. Most had been, many of them several times (Table 1).

People were then asked whether they thought most opinion polls worked for or against the best interests of the general public (Table 2). Three out of four said polls worked for the public's interests—over nine times as many as said they worked against the public's interests.

---

[1] The survey was based on face-to-face interviews with a nationwide sample of 2000 adults age 18 and up. Down to the block level the sample was a probability sample of the 48 contiguous states. At the block level, sex, age, and employed women quotas were imposed, with evening and/or weekend interviewing required to insure proper representation of employed people. Since quota controls were imposed and specific respondents were not designated, "completion rate" is an inappropriate consideration in this sample.

The questions cited in this article were asked at about the 40-minute mark in a wide-ranging 45-minute interview—after questions about international affairs, social issues, lifestyles, attitudes toward business, public radio and television listening, and an assortment of other subjects. Since nearly half of all respondents had never been interviewed before, this subject matter—and length—had to affect their attitudes toward polls, particularly with respect to the questions reported in Tables 8, 9, and 10. A similar context and length would also influence responses to the question that asked about agreement with Letter X or Y: While more than half of all respondents had been interviewed before, nevertheless this interview was their most recent experience and undoubtedly had some effect on their responses.

Abstract The Roper Organization conducted two surveys in 1985 on the public's attitudes toward polls. One was a 10-question interview and the other was a single question asking for reactions to letters both critical and laudatory of polls. Results of the two surveys, which were asked of national samples of adults, indicate that although there is skepticism among the public about polls, few people are entirely negative. However, the rise in the last 10 years in the percentage of those who are critical of polls suggests that there is cause for concern. The author makes suggestions for improving polling from the respondent's point of view.

Burns W. Roper is Chairman of The Roper Organization, Inc.

Public Opinion Quarterly Vol. 50:10–16 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-10/$2.50

Table 1. Previous Interviews

Aside from this interview, about how many other times, if any, have you been interviewed in a poll—either in person or by telephone?

Never before	41%
Once	17
Twice	16
3 to 5 times	16
6 or more times	9
Don't know	1

Table 2. Do Opinion Polls Work for the Public's Interests?

Now I'd like to ask you some questions about polls in general, based on what you know or have read or heard about them; and not just your experience in this interview. First, do you think most opinion polls work for or against the best interests of the general public?

For	75%
Against	8
Don't know	18

Table 3. Influence of Polls

How much do you think poll results affect or change things—would you say polls usually have almost no influence, or some influence, or quite a lot of influence, or have too much influence?

Almost no influence	9%
Some influence	58
Quite a lot of influence	21
Too much influence	5
Don't know	7

Table 4. Accuracy of Polls

Do you think polls are almost always accurate, usually accurate, accurate only sometimes, or accurate very seldom?

Almost always accurate	10%
Usually accurate	46
Only sometimes accurate	32
Very seldom accurate	6
Don't know	6

How much influence do poll results have in terms of affecting or changing things? Most said "some" influence. Few said either none or too much (Table 3).

Next, how accurate are polls and how honest are pollsters? More said "accurate" than said "inaccurate," but not by an overwhelming margin (Table 4).

Table 5. Honesty of Pollsters

Do you think pollsters are almost always honest about their polls, are usually honest, honest only sometimes, or honest very seldom?

Almost always honest	25%
Usually honest	51
Only sometimes honest	14
Very seldom honest	3
Don't know	7

Table 6. Representativeness of Polls

Do you think a sample of 1500 or 2000 people can accurately reflect the views of the nation's population, or that it's not possible with so few people?

Can accurately reflect views of nation	28%
Not possible with so few people	56
Don't know	15

Table 7. Are Interviewees Typical?

Do you think pollsters interview typical, representative people, or that they get mostly unusual, nontypical, or even oddball types of people?

Typical, representative people	80%
Unusual, oddball types	7
Don't know	13

Pollsters fared better on their honesty than on the accuracy of their product. Many more said "honest" than said "not honest" (Table 5).

One of the most interesting questions in the survey was the next one, which asked people whether they think a sample of 1500 or 2000 people can accurately reflect the views of the nation's population, or that it's not possible with so few people (Table 6). By a two-to-one margin people said it's not possible. In fact, exactly the same percentage said that it is not possible as had earlier said that polls were either almost always or usually accurate (see Table 4)—56 percent in both cases.

Roper asked people whether they felt pollsters interview typical, representative people, or mostly unusual, nontypical people (Table 7). Eight in 10 said typical; less than one in 10 said atypical.

People were asked how many survey respondents they think tell the truth in polls (Table 8). Eight in 10 said most or nearly all people tell the truth. Very few think that very few tell the truth.

Table 8. Do Respondents Tell the Truth?

Do you think nearly all people who are interviewed in polls tell the truth, or most people do, or only some do, or that very few in a poll tell the truth?

Nearly all	28%
Most	53
Only some	13
Very few	3
Don't know	3

Table 9. Are Polls Enjoyable?

From the point of view of the person being interviewed, do you find polls an enjoyable and satisfactory experience, or an annoying and unsatisfactory experience, or somewhere in between?

Enjoyable and satisfactory	51%
Annoying and unsatisfactory	5
Somewhere in between	42
Don't know	1

To close the survey, respondents were asked whether they find polls an enjoyable and satisfactory experience, an annoying and unsatisfactory experience, or somewhere in-between. Half said enjoyable and satisfactory (Table 9). Nearly as many said somewhere in-between—which is fine, if in fact they weren't really like the disgruntled 5 percent, but too polite or timid to say so.

The final question asked the disgruntled 5 percent why they found polls annoying and unsatisfactory, and used a card that listed seven criticisms or complaints about polls (Table 10). The most common complaint, cited by nearly two-thirds of those who find polls annoying and unsatisfactory, was that they take too much of your time. (This question was asked at about minute #40 in a 45-minute face-to-face interview.) The second most common criticism, also related to time, was that polls interrupt you at inconvenient times. Fortunately, only 1 percent of the disgruntled said that the interviewers are often unpleasant. But, of course, telling that to the interviewer who was asking the question would require nerve, or bad manners, or both on the respondent's part.

This poll raises at least two questions. One is the matter of sample bias, and the other is the honesty of people's responses to these questions. Will those people who either dislike or disapprove of polls agree

Table 10. Why Are Polls Annoying?

Which of the reasons on this card, if any, are your reasons for finding polls annoying and unsatisfactory? Any others?

	% of the Total Sample	*% of Those Who Find Interviews Annoying and Unsatisfactory*
They take too much of your time	3%	63%
They interrupt you at inconvenient times	2	39
The questions are often too personal, nosey, and intimate	1	28
The questions are often stupid or unanswerable or both	1	26
The questions are often too difficult and complicated	1	24
Too often there is some other purpose to the poll, trying to sell you something or finding out when no one in your family is home	1	22
The interviewers are often unpleasant	—[a]	1
None	—[a]	2
Don't know	—[a]	1

[a] Less than 0.5%.

to be interviewed in the same proportions as those who are not critical of polls? Logic would say that more of the critics would fail to cooperate. And will the critics who agree to be interviewed tell it like it is, or will they pull their punches and shade their answers? How many of the 42 percent who said being interviewed is somewhere between enjoyable and annoying really meant annoying and unsatisfactory?

My judgment—guess, really—would be that sample bias and shaded responses could make the foregoing results 10 points more favorable than they really should be.

In the poll just described, respondents had eight opportunities to express one or more of 11 reservations about some aspect or other of polls. Three out of four respondents interviewed did, in fact, express one or another reservation, the single most common reservation being that you can't accurately reflect the nation's opinions with a sample of 1500 to 2000 people.

In addition to this poll, Roper has asked a single question on four different occasions in which respondents were shown excerpts of two letters that were actually received from survey respondents. One of these letters was highly critical of being interviewed and the other highly laudatory about the experience. At the end of four different surveys, we asked respondents to read the two excerpts and to tell us which letter they most agreed with. In the first three askings, which covered the mid- and late 1970s, 3 percent agreed with Letter X, the critical letter. More than 90 percent on each occasion agreed with the laudatory letter, Letter Y. However, on the latest asking, which was two years ago, agreement with Letter X had doubled to 6 percent; agreement with Letter Y had dropped from 94 percent to 86 percent, and the undecided rose to its highest level ever, 8 percent.

Excerpt from Letter X:

> On last Saturday, one of your representatives came to my home and asked if she could interview me.
>
> I now regret that I cooperated and granted the interview. I feel that it was an imposition to take that much time purely for the benefit of your clients. I also feel that some of the questions asked for my opinion about things that are entirely my own business.

Excerpt from Letter Y:

> I have just given an interview to one of your interviewers and wish to thank you for the opportunity to make my opinions known.
>
> I believe you perform a vitally needed service in times like these by giving people the chance to make their opinions known—and feel that eventually it will do some good. If leaders in industry and government can know what people really think they can do a better job.

There are the data. What conclusions can be drawn from them? The question as to what these data mean would be a matter of judgment, even if there were no question about what the numbers themselves really should be. And judgment, unlike sampling error, is not subject to a mathematical measure of error.

The two studies indicate—to me, at least—that we in the opinion research field don't get a clean bill of health, but we don't face a crisis of public confidence either. The studies say there is a healthy skepticism among the American public about polls—with equal emphasis on "healthy" and "skepticism." Most people are not automatically seduced, blindly taken in, and convinced by polls. On the other hand, few appear to be irate, up in arms, or angry over them.

We *do* face *some* problems; the suggested rise in agreement with Letter X, even if that increase is not statistically significant, is cause for concern.

There are areas in which we need to clean up our act a little. The biggest justifiable complaint that the public has with polls (that are in fact polls) is the matter of interview length. Face-to-face interviews of an hour or more and telephone interviews of a half hour or more are too long unless very special arrangements are made—somewhat too long, I believe, in terms of respondent fatigue and the validity of a respondent's answers, but substantially too long in terms of imposing on a respondent without his or her express consent.

Many researchers are sometimes guilty of asking questions that are too personal—or at least too personal for many subgroups of the population ("How often do you have a daily [sic] bowel movement?").

A third criticism that can be leveled at opinion researchers—and it's the number one criticism of phony calls—is taking advantage of respondents in terms of unexpected follow-ups or attempted sales. When a person answers that her refrigerator is 15 years old, that should not subject her to unwanted sales calls from an appliance dealer or manufacturer. When a person agrees to be interviewed, that does not mean that the survey organization has license to call back on three or four additional follow-up interviews over the coming year unless the respondent has agreed to such follow-ups.

There is not a serious problem now—fortunately—but as polling gets more and more pervasive, the problem will get worse, even if our actual practices don't change one whit. The same practices in more and more polls may be what is causing the agreement with Letter X to rise.

We have to monitor and regulate our practices. What would the impact be on our major excess if our code of ethics were changed? Suppose in the responsibilities-to-survey-respondents' part of the AAPOR code researchers were required to tell survey respondents within three minutes of the start of an interview how much longer that interview would take. Not "Could I ask you a few questions?" or "This survey will take just a few minutes." But "This survey will take a total of an hour and ten minutes. Is now a convenient time to talk to you, or would it be better if I call back at another time?" If we were required to so inform survey respondents, might not interview lengths drop appreciably? Even if they did not drop appreciably, survey respondents would probably feel better about having been told up front. They would feel less like they had been had.

Trends in Attitudes Toward and Participation in Survey Research

STEPHEN SCHLEIFER

Because the survey research industry depends on the public's willingness to participate in its surveys, public goodwill must be a top priority for the nation's survey research community. To this end, Walker Research Inc.[1] has since 1974 conducted a biennial Industry Image Study, which examines the public's attitudes toward the survey research industry and some of its specific practices. Interviews are conducted with a national sample of respondents to determine their frequency of participation in polls and surveys, their opinion of the interview experience, as well as their feelings about the research industry in general. These biennial Industry Image studies have monitored trends in public opinion of survey research.

This article summarizes the findings of the 1984 Industry Image Study, which measures participation levels in surveys, attitudes toward the interview experience, and feelings about the survey research industry. (See Appendix for question wording.) The results and the trends observed from prior years provide clues on how the industry can enhance its image among the general public.

METHOD

In the week of August 20–26, 1984, 495 telephone interviews were completed[2] with a national probability sample of all U.S. telephone-

---

[1] Walker Research, Inc., is the fourteenth largest marketing research firm in the country, according to *Advertising Age* (5/23/85). Headquartered in Indianapolis, IN, Walker has eight branches nationwide and employs over 1000 people.

[2] Interviews were conducted from Walker Research's National Calling Center located in Indianapolis, IN, using its automated CRT Interviewing System. Three attempts were made on each sample household using a day-night calling pattern. The national sample was randomly generated by computer to include both listed and unlisted telephone numbers.

Abstract This study, now in its sixth wave, examines trends in consumers' willingness to be surveyed as well as attitudes toward being interviewed and toward the survey research industry in general. The results raise some important questions in terms of multiple participation and its effect on the validity of survey data. They also challenge researchers to examine their own standards and practices in order to enhance the relationship between researchers and the participants.

Stephen Schleifer is Senior Vice President of Walker Research, Inc. of Indianapolis, IN. This article is a revised version of a presentation given at the 40th Annual Conference of AAPOR at McAfee, NJ, May 18, 1985.

Public Opinion Quarterly Vol. 50:17–26 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-17/$2.50

Table 1. Limitation of Findings/Refusal Summary[a]

Representative Sample	*1978*	*1980*	*1982*	*1984*
Completed interviews (*N*)	500	500	499	495
% Complete	51%	59%	54%	54%
Total refused	477	346	432	415
% Refused	49%	41%	46%	46%
At introduction	22%	16%	30%	40%
During interview	27%	25%	16%	6%

[a] The findings of the 1984 Industry Image Study are limited by the normal statistical error inherent in any sampling process. Because the research was conducted by telephone, it is not necessarily representative of nontelephone-owning households. Further, increased potential for nonresponse bias exists in the findings due to the relatively high level of respondent refusals (as has been observed in past studies).

owning households. Of these, 115 people (23 percent) reported participation in at least one survey in the past 12 months. To increase the number of these past year participants, additional interviews were completed. Therefore, the "representative sample" includes 495 respondents, including 115 who claimed survey participation in the past year. The "past year participant sample" includes the 115 respondents chosen at random, *plus* 86 obtained on an over-quota basis (for a total of 201 people).

The first two waves of the Industry Image Study (in 1974 and 1976) included interviews from only the top 30 markets. In 1978, the sampling plan was modified to a national probability design. The 1984 results can therefore be directly compared to the 1978, 1980, and 1982 surveys. The response rates achieved since 1978 are shown in Table 1.

RESULTS

Survey Participation. Annual studies conducted by industry groups such as the Council of American Survey Research Organizations (CASRO) and the American Marketing Association (AMA) estimate participation levels in surveys based on figures provided by research organizations and manufacturers. By contrast, this study's projections emanate from the reports of survey participants themselves. According to the 1984 results, overall participation in polls and surveys has remained constant since 1982, which was down somewhat from 1980. As illustrated in Figure 1, 54 percent of the representative sample said they had been interviewed at some point but not in the past year, while 23 percent said they had been interviewed within the last 12 months. Identical results were obtained in 1982. By contrast, 1978 levels were

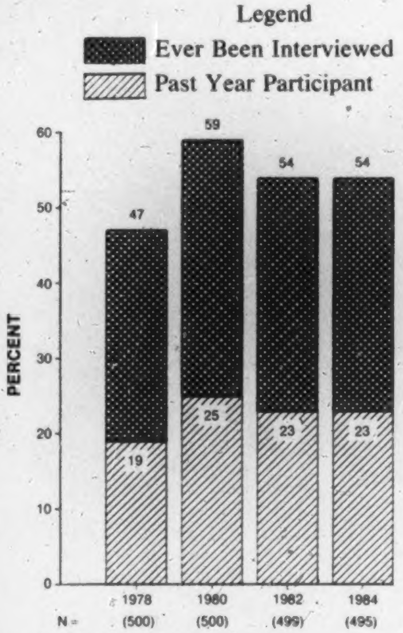

Figure 1. Participation Levels Among the Representative Sample

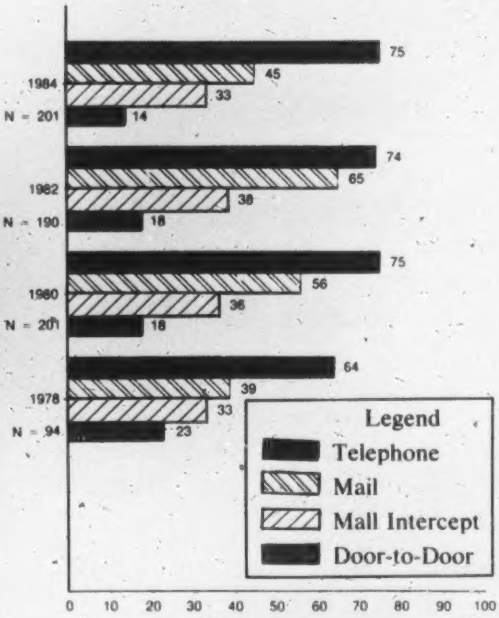

Figure 2. Participation by Type of Interview: Past Year Participation

significantly lower[3] than in any of the following years. These findings coincide with the data published annually in *Advertising Age* showing the real growth of survey research slowing from the explosive expansion observed in the late 1970s.

The type of interviews respondents participated in over the past year remains basically unchanged from prior measurements, as shown in Figure 2. According to 1984 past year participants, telephone continues as the dominant form of data collection (75 percent) followed by mail (45 percent), mall intercept (33 percent), and door-to-door (14 percent) surveys. After experiencing rather dramatic increases in recent years, mail surveys showed a significant *decrease* in 1984. Door-to-door interviewing also continues to decline as a viable method.

With this significant amount of survey activity, concern has been voiced about the "freshness" of those sampled. Are researchers reaching new participants each time, or are they interviewing the same respondents survey after survey? Prior years' results suggested that multiple participation is occurring and might be growing. To quantify this hypothesis, 1984 respondents who said they were interviewed in the

---

[3] Throughout this article, the term *significant difference* reflects a significant change using an independent *t*-test at the 95 percent confidence level.

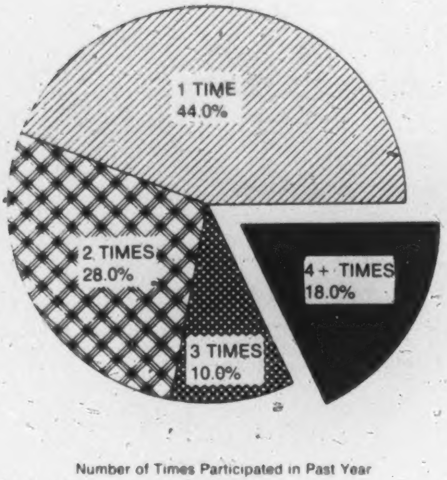

Figure 3. Repeat Participation in Past Year

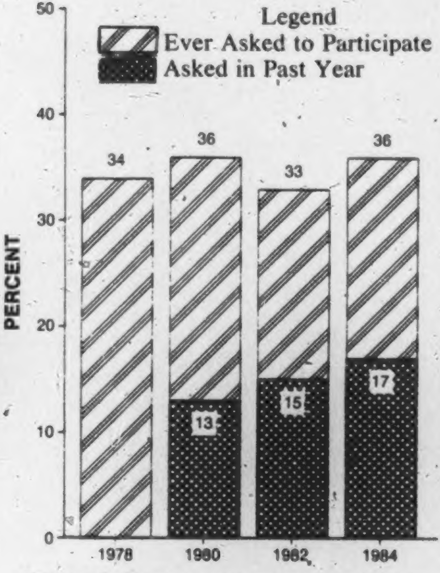

Figure 4. Exposure to a Disguised Sales Pitch

Table 2. Multiple Participation in Single Method

	1978	*1980*	*1982*	*1984*
Base: past year participants (*N*)	94	201	198	201
	%	%	%	%
Telephone	51	56	43	50
Door-to-door	22	42	30	43
Mall intercept	25	43	48	26
Mail	43	57	58	47

past year were asked how often. Figure 3 shows the majority of past year participants (56 percent) said two or more times, while 44 percent said they were surveyed just once over the past 12 months. Of concern is that 18 percent of the past year participants said they had been interviewed four or more times.

Further questioning into repeat participation by method type revealed a reversal in 1984 from the trends observed in 1982, as shown in Table 2. After a significant decrease in 1982 (from 1980), repeat participation in telephone interviews rebounded somewhat in 1984, returning to the level first observed in 1978. As for mail surveys, this research reveals a significant decrease in multiple participation comparing 1984 data to 1980 and 1982. Another decrease over the past year is seen for mall intercept studies. After a steady rise in reported repeat participa-

Table 3. Demographic Profile of Past Year Participants

	1984 Participants	*1984 Nonparticipants*
Base: Total respondents (*N*)	201	380
Education	%	%
High school grad or less	46	60
Some college or more	52	36
Income		
Under $20,000	26	36
$20,000 plus	51	41
Age		
Under 35	43	27
35 plus	54	67
Sex		
Male	21	46
Female	79	54

tion, the significant decrease found in 1984 will be encouraging to some, and may perhaps reflect an increased emphasis to screen out frequently intercepted respondents.

It is difficult to draw definitive conclusions from these data on repeat participation. However, they suggest the need to address some important questions:

> Are there segments of consumers more willing to cooperate, who consequently tend to be overrepresented in surveys?
>
> How much is "too much?" At what level of repeat participation should researchers become concerned?
>
> To what extent does more frequent participation in polls and surveys create "professional" respondents rather than typical consumers? What is the effect of multiple participation on the validity of the data collected?
>
> Should previous participation become a standard screening question for polls and surveys? Is it worth the expense to include only "fresh" respondents each time?

These and other questions should be addressed by those of us engaged in survey research. While we can only speculate about the characteristics of people who steadfastly refuse to participate in surveys, the Industry Image Study does reveal some things about those who do. Table 3 compares the demographic characteristics of past year participants with those who had not been surveyed over the same period of time. Past year participants in 1984 were significantly more likely than nonparticipants to be under 35 (43 percent versus 27 percent), to have completed at least some college (53 percent versus 36 percent), to have income over $20,000 (51 percent versus 41 percent), and to be female (79 percent versus 54 percent). These demographic differences are consistent with those reported in previous waves of this research.

The increasingly targeted nature of marketing researchers' samples and the entire industry's continued growth may in part explain the high levels of repeat participation in surveys. While public opinion surveys typically sample a cross-section of the general public, marketers' growing use of market segmentation requires that they look for specific population groups which are thought to be the best consumers for their goods or services. This study suggests that some of these groups—the young, better educated, higher income, females—are of particular interest to many, and that they are oversurveyed. If the trend continues, it may become more difficult to interview these popular segments, and both refusal rates and repeat participation may rise.

False Surveys. A factor contributing to the public's attitudes toward the research industry is false surveys—selling efforts under the guise of survey research. According to the 1984 Industry Image Study, reports of such false surveys remain at relatively high levels. As shown in Figure 4, over one-third (36 percent) of those in the representative sample said they had at some time been asked to participate in an alleged poll or survey which turned out to be a sales pitch. The trend is moving upward to the point where 17 percent of 1984 past year participants said they were subject to one or more false surveys. These disguised sales pitches harm survey researchers' reputations, raise the public's suspicions, and contribute to increased refusal rates.

Attitudes Toward Being Interviewed. Based on most industry projections, the level of survey activity will continue to grow as the demand for feedback from the marketplace accelerates. Researchers' ability to meet this demand depends on the public's attitudes about being surveyed. What, therefore, does the public think of survey research which asks much of them, and may appear to give little or nothing in return?

The 1984 results show a downward trend in those past year participants (79 percent) who rated their last interview as a "pleasant" experience (see Figure 5). In addition, increasing numbers said that the interview was too long (23 percent) and contained what participants felt were confusing (16 percent) and overly personal (14 percent) questions. As shown in Figure 5, the percentage of past year participants who said their last interview was an "unpleasant" experience rose significantly from 6 percent in 1982 to 12 percent in 1984. While the other shifts from 1982 to 1984 in Figure 5 are not statistically significant, the directional implications are clear and consistent. The data do suggest, however, that researchers can exercise some control

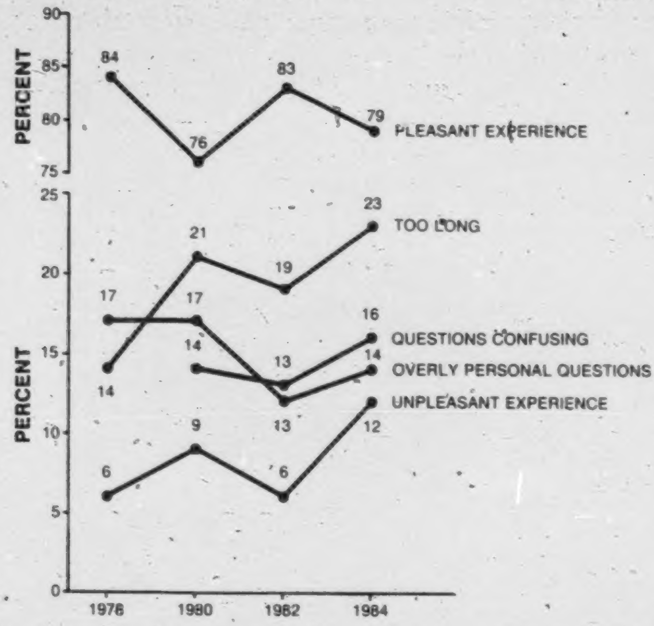

Figure 5. Attitudes Toward Being Interviewed: Last Interview Evaluation

over respondents' attitudes. Shorter interviews, clearly worded questions, and the elimination of unnecessary personal or sensitive items would help. Determining what is an optimum interview length or which questions are perceived as too personal are issues for further study.

There is reason to believe, however, that the element most critical in shaping respondents' attitudes is the interviewer. For the first time in 1984, past year participants were asked why they had evaluated their last interview as either "pleasant" or "unpleasant." The results showed that the personality or attitude of the interviewer was cited most frequently (47 percent) by those who said it was a pleasant experience. Other potential influencing factors such as questionnaire design (20 percent), the study's particular topic (8 percent), or whether the interview was conducted at a convenient time (3 percent) were mentioned less frequently. Only 1 percent (3 respondents) said the experience was unpleasant. These findings have obvious implications for interviewer and supervisor training, practices, and procedures.

Public Attitudes Toward the Industry. The public's overall attitudes toward the survey research industry are somewhat more positive than those expressed about their actual interview experience. These attitudes were assessed by asking respondents to agree or disagree with a series of belief statements. As shown in Table 4, the majority (81 percent) of respondents feel the research industry serves a useful purpose,

Table 4. Image of the Industry

	1978	*1980*	*1982*	*1984*
Base: representative sample (*N*)	500	500	499	495
Positive image attributes	%	%	%	%
The research industry serves a useful purpose	85	85	84	81
Polls and research surveys are used to help manufacturers produce better products	81	80	80	82
Answering questions in polls or surveys is an interesting experience	70	64	66	60
Answering questions in polls or research surveys is in my own best interest	72	70	68	60
Survey research firms maintain the confidentiality of answers	—	60	61	54
Negative image attributes				
Some questions asked in polls or research surveys are too personal	42	47	49	51
The information obtained in polls or research surveys helps manufacturers sell consumers products they don't want or need	38	45	43	44
The term *poll or research survey* is used to disguise a sales pitch	43	38	38	39
Polls cr research surveys are an invasion of privacy	31	25	26	28
Answering questions in polls or research surveys is a waste of time	17	18	21	22
The true purpose of some surveys is not disclosed	—	68	65	61

and that surveys are used to help manufacturers develop better products (82 percent). On the other hand, only 54 percent of the 1984 representative sample believes survey research firms maintain the confidentiality of answers, compared to 61 percent who held this belief in 1982. This finding is particularly disturbing because researchers must have the public's trust to gain access and obtain valid data.

Overall, based on the answers to these industry-related questions, the signals being sent to researchers are mixed. On the one hand, some attitudes suggest the public doesn't object to surveys and polls *in principle*. On the other hand, many people remain suspicious of the industry's methods and motives.

CONCLUSIONS

The data obtained in this ongoing study of attitudes towards the survey research industry are generally encouraging, although several problem areas do exist. These problems must be addressed by *all* segments of the research industry in order to maintain and improve the quality of data collected via polls and surveys. Because the industry depends on public cooperation for its livelihood, efforts must continue to solidify the relationship between researchers and potential respondents. The results of this survey indicate that while the public is supportive of the research industry in general, some of its perceived practices appear to be eroding these positive feelings. Interviewers who are poorly trained and supervised, interviews that are too long and contain overly personal questions, and repeated requests to participate in surveys appear to be among the causes. To reverse this, all those engaged in survey research must continually evaluate their own practices from the *respondent's* point of view, and carry out their investigations in a way that will not harm the industry's image.

APPENDIX: 1984 INDUSTRY IMAGE STUDY

1. Thinking back, have you, yourself, ever participated in a survey before, or not?

2. Have you, yourself, been interviewed for a survey in the past year, or not?

3a. Thinking about the past year, how many times have you, yourself, been interviewed for a survey?

3b. And now, thinking *just of the past year*, about how many times have you, yourself, been interviewed for a survey. . . By telephone? In person at home? In person at a store or shopping center? By a questionnaire sent to you in the mail?

4a. Now I'd like you to think of the *last* time that you participated in a poll or research survey. Was that particular interview by telephone, in person at home, in person at a store or shopping center, by a questionnaire sent in the mail, or by some other method?

4b. Still thinking of the *last* time you were interviewed, were any of the questions worded in a confusing manner, or not?

4c. Were you asked any questions that you thought were overly personal during that last survey, or not?

4d. What did you find overly personal about those questions?

4e. About how long, in minutes, did that last survey take?

4f. In your opinion, did the survey last too long, or not?

4g. And, still thinking of the last time you were interviewed, was your *overall* impression of the interview pleasant or unpleasant?

4h. You said your overall impression of the interview was *pleasant*. Why did you say that? What other reasons? Why else?

4i. You said that your overall impression of the interview was unpleasant. Why did you say that? What other reasons? Why else?

4j. What was the topic or subject of that last interview?

5. Now, I'm going to read a list of statements which some of the people we've talked with have said about polls or research surveys. As I read each statement, please tell me whether you agree or disagree with the statement.

STATEMENT

The survey research industry serves a useful purpose.

Answering questions in polls or research surveys is in my own best interest.

Answering questions in polls or research surveys is a waste of time.

The term "poll or research survey" is used to disguise a sales pitch.

The information obtained in polls or research surveys helps manufacturers sell consumers products they don't need or want.

Some questions asked in polls or research surveys are too personal.

Polls or research surveys are used to help manufacturers produce better products for consumers.

Polls or research surveys are an invasion of privacy.

Answering questions in polls or research surveys is an interesting experience.

Poll or research survey companies should not be allowed to interview people in shopping centers.

Survey research firms maintain the confidentiality of people's answers.

Research surveys give people an opportunity to provide feedback to manufacturers and retailers about the products or services they provide.

The true purpose of some surveys is not disclosed to the people who participate.

6a. In the past year, have you refused to participate in what appeared to you to be a legitimate poll or research survey, or not?

6b. Have you *ever* been asked to answer questions in a poll or research survey that turned out to be a sales pitch for a product or a service?

6c. Have you been asked to answer questions in a poll or research survey which turned out to be a sales pitch for a product or service *in the past year*, or not?

6d. What types of products or services did the sales pitch deal with?

7a. Surveys are conducted on a wide variety of subjects or topics. I would like to read a short list of different topics and ask you how likely you would be to participate in such a survey.

The first/next topic is (*TOPIC*). How likely would you be to participate in a survey that asked (*STUDY*)? Would you say you would participate, might or might not participate, or would not participate?

TOPIC (STUDY)

Automobiles (for your reaction to new automobile designs).

Political candidates (who you plan to vote for in the next election).

Food products (for you to try a food product and get your opinion of it).

Advertising (for you to recall what advertisements you have seen recently and what you think of them.

Television (you to describe your normal television viewing habits).

Telephone company (you to evaluate the service you get from the telephone company).

Personal hygiene products (for you to evaluate a specific brand of personal hygiene product).

7b. The telephone is used for a variety of purposes, including surveys as well as for other purposes. I would like to read a short list of the ways the telephone is sometimes used and ask you how acceptable each would be to you if you were called for such a purpose.

One/Another way the telephone is used is (*PURPOSE*). How acceptable would such a call be to you? Would you say it would be very acceptable, somewhat acceptable, or not at all acceptable?

PURPOSE

Calls from the blood bank asking you to donate blood.

Charitable organizations asking for contributions.

The selling of insurance.

Handicapped people selling household products.

Organizations such as Goodwill or AmVets asking you to donate used goods.

Real estate agents looking for business.

Companies selling home improvement products such as aluminum siding.

Political organizations asking you to vote for specific candidates.

Surveys or research companies asking you for your opinion on a product or service.

Now I have just a few last questions to help us group your answers with those of others who might be somewhat like you.

[A–D are demographic.]

E. Finally, how long would you estimate that this interview lasted?

F. Do you consider that to be "too long" or not?

Surveys on Surveys: Limitations and Potentialities

JOHN GOYDER

A "survey on surveys" questions respondents about attitudes and past actions concerning surveys, the pioneering exemplar being Sjoberg (1955).[1] Several aspects of such data have interested survey researchers: The extent of bogus surveying has generated sustained attention (Arnold, 1964; Baxter, 1964; Biel, 1967; Allen and Colfax, 1968; Rugg, 1972), and questions about preferred form of contact appeared in investigations of the acceptance of telephone surveying (Rogers, 1976; Groves, 1979:193). Attention to sensitive question topics was reported in Singer (1984), measures of "response burden" have been fielded by Sharp and Frankel (1983), and an emphasis on recording motives for refusal appears in studies such as DeMaio (1980). Day (1975:464) discussed the legitimacy the public attributes to surveying. Schwartz (1964) collected data on most of the above themes, but perhaps the most comprehensive survey on surveys is a U.S. Census Bureau data set reported in National Research Council (1979).

---

[1] I have adapted the term from Sjoberg's (1955) "Questionnaire on Questionnaires."

Abstract Data from two surveys on surveys in Waterloo, Ontario were used to assess attitudes toward the survey. Questions in the two surveys inquired into motives for refusing/consenting to interviews; attitudes toward past experience with surveys; likes/dislikes of form of contact, question format, and topic; perception of social pressure to respond to surveys; views on the social legitimacy of polling and on legislative controls for surveying. The surveys also collected response histories. Analysis of the attitudes began with a factor analysis of 12 items, which identified four factors descending from the general to the particular. It was found that attitudes are related to exposure to surveying, i.e., the more times a person reports requests for survey cooperation, the more unfavorable is his or her attitude toward the method. On the whole, however, the author finds that the evidence presented belies pessimistic views about surveys on surveys.

John Goyder is Associate Professor, Department of Sociology, University of Waterloo, Ontario, Canada. This article is a revision of a paper presented at the 40th Annual AAPOR Conference, May 16–19, 1985. Research assistants on the record-linking portion of the study were Joti Sekhon, Christine Hutchinson, Ella Haley, and Joan Lyons. Twenty-one students enrolled in Sociology 282 in Winter 1982 collected the 1982 survey data. Patricia Anderson, Helen Chapman, and Joan Lyons conducted telephone interviews for the 1985 replication. Marnie Goyder was coder on the replication. The work reported herein has been supported by the University of Waterloo Faculty of Arts Resource Fund, the UW-SSHRC Small Grants Subcommittee, and by an SSHRC Leave Fellowship. The hospitality of the Social and Political Sciences Committee, Cambridge University, in granting the author Visiting Scholar privileges in 1985–86, greatly assisted the completion of the manuscript.

Public Opinion Quarterly Vol. 50:27–41 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-27/$2.50

A new wave of survey on surveys data, reported in this issue of *POQ*, makes a reassessment of the limitations and potentialities of the data timely. By some criteria, the pool of survey data on surveying has become impressive. Where earlier research was often conducted on a local scale, and on a small budget (e.g., Sjoberg, 1955), many of the recent data derive from premier market research firms such as Gallup and Roper. The new studies thus often read public opinion on surveying at the national level, with replication at least of main themes, if not of specific questions, across survey houses.

On the debit side, however, the mounting quality of survey on surveys data serves only to sharpen more fundamental reservations about such investigations. Receiving much attention in the recent reports is the problem of "professional respondents" who, with nonrespondents by definition absent from tabulations, bias results in favor of public approval of surveying.[2] Another concern, potentially the more damaging, is the likelihood that public attitudes toward surveying amount to no more than unstable surface opinions, rather than definable attitudes in any psychological sense (Rokeach, 1968:455).[3]

The present article reports material that addresses these two aspects of the validity of surveys on surveys. The data are confined to one municipality but, as we shall argue in the following description of research design, they complement the national studies by addressing validity threats explicitly.

DESIGN

The epistemological limitation to surveys on surveys is self-evident; employing an instrument to measure its own performance is immediately contradictory. To assess attitudes toward the survey via attitude surveys is not unlike seeking to comprehend the mechanics of the camera solely from photographs. The contradiction becomes at least partially resolved, however, when survey data are placed alongside external sources of evidence, and in the present case this design was achieved by record linking.

Two surveys on surveys were collected in Waterloo, Ontario, Canada. The first, fielded in March 1982, employed student interviewers who attempted contact at the door with 140 names randomly (cluster) sampled from the city directory.[4] In April 1985 the survey was replicated, with professional telephone interviewers attempting contact with another 140 listings, drawn by the same method as in 1982. For all sample listings, respondents and nonrespondents alike, a data file was compiled and record-linked with the survey data. From municipal taxation records it was possible to learn property assessment, age, citizenship, and religion (Protestant–Catholic). The city directory added marital status, labor force standing, gender, some details on household composition, type of dwelling, ownership or rental of dwelling, occupation, migration history (from back issues), and a code identifying householders who refused to be enumerated for the directory. Since only a percent or two of the population refuse the directory, and because such refusals offered valuable leverage to the design, the 1985 sample was stratified in the proportion one-third city directory refusers to two-thirds random cross-section.

---

[2] "Melvin D. Field, the head of Field Research of San Francisco, said he felt that the studies were flawed because those people who answer questions in a poll are inherently more sympathetic to polls than those who refuse to be interviewed." (*New York Times*, 1985:B7).

[3] "Our image may only be as good as our last election." Andrew Kohut, President of Gallup, quoted in *The New York Times* (1985:B7).

Question design in the 1985 survey was identical to the 1982 schedule, excepting the deletion of a few of the less useful 1982 questions in the interests of streamlining the telephone enumeration. In final (1985) form, the interview schedule required some 15 minutes to administer and inquired into: attitudes toward past experiences with surveys; motive for refusing and consenting to interviews; likes and dislikes concerning form of contact, question format, and question topic; perception of social pressure to respond to surveys; general views on the social legitimacy of opinion polling; and attitudes on legislative controls on surveying. The attitude questions were most directly inspired by Sjoberg (1955), from whom several wordings were reproduced, but most of the topics investigated have appeared in the other literature cited above.

In addition to the attitude questions, the surveys collected response histories, reconstructing past contacts by surveyors, and eliciting details on the ensuing response decisions. We regard the response history as partially external to the attitude questions. Although collected within the same instrument as the attitudes, and thus subject to the same nonresponse bias and other errors, the response history constitutes a behavioral variable against which the attitude items can be validated.

---

[4] Time and resources permitted only two systematic calls, with the occasional third call to fulfill an appointment. For the 50 cases in which contact was not achieved, the survey instrument was mailed out. It had been designed to be self-administerable, since it was known that multiple callbacks at the door would be impossible. Fuller fieldwork design particulars for the 1982 survey appear in Goyder (1985a).

The overall case base for the analysis is small (280 sample listings, totaled over the two surveys, with 172 responses to the surveys), making .05 significance hard to achieve. Where key computations are at stake below, successful reproduction across samples of the direction of a relationship is viewed as supporting evidence given a marginal ($p < .1 > .05$) result on the formal significance test.

The article begins with the problem of "professional respondents," exploiting the record-link file for a nonresponse analysis of the Waterloo surveys. An assessment of the validity of attitudinal data about surveys is then attempted, using two criteria. The stability of factor analysis results for the attitude items from the two Waterloo surveys is tested by comparing loadings from the two data sets, and then the attitude constructs from factor analysis are validated against the history of response behavior.

RESPONSE BIAS IN THE SURVEYS ON SURVEYS

Fifteen of the total of 280 sample addresses drawn for the 1982 and 1985 surveys proved vacant and are removed from further analysis. Response on the remaining case base of nonvacant households was 65 percent in each sample. The telephone contacts involved up to five calls, a plan consistent with available resources and deemed likely to yield a response in the sixties. It seemed consistent with accurate replication to engineer equivalent response on each sample. Tabulations for analysis comparing respondents against nonrespondents appear in Table 1.

Logistic regression is employed, since response is binary. Thus, the dependent variable is the natural log of $p/(1 - p)$, where p is probability of response, and the regression coefficients entered into Table 1 are computed on the maximum likelihood criterion (Hanushek and Jackson, 1977:200–3).

Entered individually, as zero-order effects, seven sample listing characteristics passed .05 or .01 significance. (Given the small case base, we adopt in Tables 1 and 4 the convention of describing $p < .05$ as "significant," $p < .1 > .05$ as "marginally significant," $p < .5 > .1$ as nonsignificant but worth entering in the tables, as possibly reproducible in a larger data set, and $p > .5$ as so certainly due to chance as to not justify reporting). Since zero-order effects among sociodemographic characteristics are likely to be reexpressions of one another, a multivariate model was next computed, in which the independent variables were successively entered alongside occupational SES. SES is operationalized through two indicators: the Blishen scale (Blishen and McRoberts, 1975), a Canadian equivalent of the Duncan scale, and a dummy variable identifying cases where occupation was either uncod-

Table 1. Nonresponse Analysis, Waterloo Surveys

	Logistic Regression Coefficients (Dependent Variable = Proportion Responding)		
		Multivariate	
Predictors	*Zero-order*	*Net of SES*	*Net of SES and Noncontact*
Intercept	—[a]	—[a]	—[a]
SES	.039***	—[a]	—[a]
No data, SES	−.527**	—[a]	—[a]
Age	−.019**	−.018*	−.027***
City directory refuser	−.339		
Religion	−.261*	−.275*	−.285*
Marital status	.854***	.539*	.358
Labor force status	.343		
Head of household	−.429		
Family type	.689		
Gender	†		
Own or rent	.532**	.268	†
Property value	.0002***	.0001	.0001
Citizenship	−1.117		
Residential stability	†		
Apartment dweller	−.839***	−.663**	−.296

[a] Not entered—differs for each computation.

* $p < .1 > .05$

** $p < .05 > .01$

*** $p < .01$

No star, $p < .5 > .1$

† Not entered, since $p > .5$

Scoring notes: SES, No data, SES: see text/Age: years/City Directory Refuser: 0 = no refusal, 1 = refusal/Religion: 0 = Protestant, 1 = Catholic, but Protestant school supporter, 3 = Catholic, Catholic school supporter/Marital Status: 0 = single (incl. divorced; widowed), 1 = married, spouse present/Labor Force: 0 = not in labor force, 1 = in labor force, incl. students/Headship: 0 = not head, 1 = head/Family Type: 0 = one generation listed, 1 = multiple generation/Gender: 0 = female, 1 = male/Own or Rent: 0 = renter, 1 = owner/Property Value, in dollars/Citizenship: 0 = alien, 1 = citizen/Stability: 0 = different listing in city directory two years prior to survey, 1 = same listing/Apartment: 0 = individual entry dwelling, 1 = apartment with common entry.

$N = 265$, with means inserted for missing data.

able or unknown.[5] The logic behind preentering occupational SES is that socioeconomic status constitutes the most enduring of all themes in studies of survey response bias (e.g., references in Goudy, 1977). No causal assumption is intended; the interest in the present analysis simply lay in identifying response biases net of the ubiquitous "middle class" bias encountered in virtually every survey. Net of SES, response remains significantly (.05) lower among apartment dwellers and marginally (.1) lower among Roman Catholics, the unmarried, and the elderly.

---

[5] Thirty-three percent of cases lacked an SES code, the largeness of the figure being partly due to the oversampling of city directory refusers. Entries in city directories are hard to code onto a fine-grained scale, since the reports often simply identify a firm. That problem was attacked by checking back entries of the directory. Often, the entries from several years summed to sufficient information for an accurate code. A bias accumulates from the back-checking because marginals to the local labor force, as well as recent in-migrants, become underrepresented. The dummy category was created to offset that bias.

In the final computation in Table 1, a dummy variable identifying confirmed contact was included in the model. The dummy virtually redefines response rate into refusal rate, since that part of response attributable to noncontact is partialed away. Apartment dwelling now drops from the model, confirming the intuition that low response from apartment buildings is attributable to the difficulty surveyors encounter in gaining access. Marital status, marginally significant in predicting total response, also slips below the $p = .1$ level. Again, the conclusion can be that response among marrieds is high for accessibility reasons (spouses suggest suitable moments to call back, etc.). Old age, however, gains strength as a correlate of low response and now surpasses .05 significance. One concludes that surveyors gain access to the young with difficulty, and this problem of access masks a relatively low refusal rate. Refusal on the city directory, the variable stressed earlier as tapping externally defined survey resistors, does not emerge as an important factor in the model of nonresponse. The effect takes the expected direction of city directory refusers tending to decline the Waterloo surveys, but falls well short of even .1 significance. The directory refusers receive further analysis below, but in the next section attention moves to the factor structure of attitudes toward surveys.

ATTITUDES ABOUT SURVEYS, AND SURVEY RESPONSE BEHAVIOR

Six questions from the Waterloo surveys were germane to constructing a general attitudes-toward-surveys variable: question 9, asking about objections to being asked questions about various topics; question 15, entailing eight agree-disagree rankings of statements about opinion polling, over a five-category scale; question 16, regarding city control of surveying; question 17, asking "should governments be the only ones allowed to conduct surveys;" question 19, inquiring whether the census should be voluntary or compulsory; and finally question 20, a Likert scale ascertaining favorability toward "the City of Waterloo conducting a regular poll on issues facing the city."

Details on question wording, together with descriptive statistics, appear in Table 2. A fuller descriptive account of the reading of public opinion toward surveying appears elsewhere (Goyder, 1985b), and for present purposes we shall simply note that, like other such informants

Table 2. Factor Analysis Results, for Respondents on Waterloo Surveys

					Factor Loadings			
Item	*Range*	*Mean*	*s.d.*	*High Score =*	*1*	*2*	*3*	*4*
For each of the following, could you say if you would object strongly, a little, or not at all, about being asked questions in a survey. (Score is total over 7 items, described in text.)	7–21	16.0	3.4	No objection	.152	.176	−.055	.279
Could you tell me whether you would strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, or strongly disagree, with the following statements about political opinion polls:								
Public opinion polls make the country more democratic.	1–5	3.2	1.2	Disagree	.163	.007	.486	−.197
The biggest problem with opinion polls is the way the media distort the meaning of results.	1–5	2.1	1.1	Disagree	−.044	.269	.163	.113
Most opinion surveyors have faked their results on one occasion or another.	1–5	3.0	1.1	Disagree	.076	.730	−.114	−.007
Government pays too much attention to opinion polls.	1–5	2.8	1.2	Disagree	.110	.058	−.110	.535
Opinion surveyors interview too few people for results to be accurate for the whole population.	1–5	2.4	1.2	Disagree	.130	.364	.046	.119
Opinion surveyors are just out to make money.	1–5	3.7	1.0	Disagree	.480	.204	−.077	.128
Opinion polls are an invasion of privacy.	1–5	3.5	1.2	Disagree	.601	.075	−.105	.047
Opinion polls do no harm.	1–5	2.8	1.2	Disagree	−.328	−.189	.499	.068
Do you think the city should set up controls to restrict the interviewing of people? Should governments be the only ones allowed to conduct surveys? (Score is summation of the 2 questions.)	2–4	3.8	0.5	No	.419	−.051	.160	.063
Some people would say that the census, which Statistics Canada collects, should be voluntary but others feel it should remain compulsory, as it now is. Would you say the census should be voluntary or compulsory?	1–2	1.8	0.4	Compulsory	−.022	.054	.213	.213
Would you be very much in favor, somewhat in favor, neutral, somewhat not in favor, or very much not in favor of the City of Waterloo conducting a regular poll on issues facing the city?	1–5	2.2	1.2	Not in favor	−.009	.063	.390	−.038

(Kohut, 1986; Roper, 1986), the Waterloo respondents return a mixed verdict on surveying. Means tend to be centered in the middle of distributions, and standard deviations are relatively high. Only on the two questions about government control of surveying is there near-unanimity: the majority of respondents in both samples did not think the city should control interviewing, and disagreed that only governments should be allowed to conduct surveys. Responses for these two questions, both skewed binary variables, were summed to give a statistically better behaved index of government regulation sentiment. Topics (money, politics, religion, family life, occupation, sex, environment) in the battery to do with objectionable question domains also were summed, to provide a generalized index of level of objection to being asked survey questions.

The righthand portion of Table 2 describes factor analysis loadings for the 12 attitude items. The computational method is principal axis factoring, and varimax rotation, using SPSS-X. Missing data seldom exceed 1 percent in the Waterloo surveys, and were handled by substituting means.

The analysis identified four factors and, in characteristic pattern, the factors tend to descend from the general to the particular. Thus, factor 1 describes those who agree that surveys invade privacy, and that "opinion surveyors are just out to make money," while disagreeing with the statement "opinion polls do no harm." Factor 1 also incorporates the feeling that government should control surveying (all loadings .3 or greater). Factor 2 exhibits high (defined as .3) loadings on just two indicators: agreement that "most opinion surveyors have faked their results on one occasion or another" and that "opinion surveyors interview too few people for results to be accurate for the whole population." The two items suggest a "methodological factor" describing disbelief in the accuracy of polls, rather than dislike of surveying itself. Polling as legitimate when confined to the political sphere seems to be captured in factor 3, which loads strongly on "public opinion polls make the country more democratic," and favorability to the notion of a regular civic poll, along with agreement that "opinion polls do no harm." Factor 4 is a one-item factor, loading .535 on the feeling that "government pays too much attention to opinion polls" and less than .3 on all other items.

The presence of replicate data from 1982 and 1985 makes a stability check on the above interpretations possible. Correlating the factor loadings for 1982 and 1985 (random portion), a method suggested in Harman (1976:343), yields $r = .78$ for factor 1 and $r = .51$ for all four loadings. We gain confidence in the first factor, and become aware that the details of the fourfold structure are vulnerable to instability.

The attitude factors were next entered, along with other variables

Table 3. Response History, Waterloo Survey Respondents

	Yes		No		Total	
Item	*N*	*(%)*	*N*	*(%)*	*N*	*(%)*
To begin with, has an interviewer ever before called at the door to do a survey—except for the census?	62	(36)	110	(64)	172	(100)
IF YES—The last time an interviewer called, did you consent to be interviewed?	53	(85)	9	(15)	62	(100)
Have you ever had a questionnaire arrive in the mail?	125	(73)	47	(27)	172	(100)
IF YES—The last time you received a mail questionnaire, did you fill it out and return it?	80	(64)	45	(36)	125	(100)
Have you ever before been telephoned to be surveyed?	108	(63)	64	(37)	172	(100)
IF YES—The last time you were telephoned for a survey, did you consent to be interviewed?	65	(60)	43	(40)	108	(100)

described below, as predictors of survey response history. The response history differentiated by form of contact, first inquiring, "Has an interviewer ever before called at the door to do a survey—except for the census?" Those reporting a request for a personal interview were then asked, "The last time an interviewer called, did you consent to be interviewed?" The same sequence then ensued for mailed questionnaires and for telephone surveys. A summary of the response history tabulations appears in Table 3. Thirty-six percent of respondents recalled a previous request for an interview at the door, against 73 percent who remembered a mailed questionnaire and 63 percent reporting a telephone contact.[6] The selection bias involved in working with respondents on a survey on surveys prohibits the drawing of any definitive conclusions about the reported response behavior on the different forms of contact. From independent knowledge (Groves and Kahn, 1979:63–76), however, it is known that telephone surveys tend to achieve lower response than personal interviews. That reality is accurately mirrored in the 85 percent responding to the last request for a personal interview, against 60 percent recalling having answered the last telephone survey. The 64 percent response figure for mailed questionnaires is plausible, assuming reasonably vigorous researcher input (Heberlein and Baumgartner, 1978).

Given that refusal probabilities correlate with form of contact, it would be useful to analyze predictors of response for each contact type separately. While we have in unreported analysis examined each component of the history (ensuring, for example, that the sign for the correlation of response history with attitude is consistent across all the tests) it is hard to feel confident about the minutiae of so detailed an investigation of so slim a data base.[7] For personal interview response, after all, the base shrinks to a mere 62 cases. Preferable was the construction of a variable summarizing the response history. Almost half the sample, 77 cases or 45 percent had, judging from their reports, refused *either* the last request for a personal interview, the last mailed questionnaire, or the last telephone contact. This information was entered into a binary variable, contrasting the refusers against the 79 cases who had never refused a survey. For comparability with Table 1, refusers are coded 0, no refusals 1. The number of requests was retained as an independent variable to be entered into analysis later. The 16 cases unable to recall a request for any of the three types of survey described in the response history were deleted.

---

[6] Forty-six percent (random sample) had never been asked for either a personal interview or a telephone survey. That figure compares closely with the 41 percent reported for a spring 1985 survey by Roper, using a different response history question (Roper Organization, 1985:17).

Logistic regression analysis of the response history summary variable appears in Table 4. Since logistic regression results convey limited face value meaning, the emphasis centers upon examining how coefficients for predictors of response alter under different modeling combinations (detected by reading across rows). Starting point, as in Table 1, is the zero-order case where predictors are entered singly. All the attitude variables correlate in the expected direction with response (unfavorable attitudes correspond to high probability of a refusal), with factor 1 exceeding the .05 level of significance, and factor 2 attaining marginal ($p < .1$) significance. Since this result corroborates our earlier judgment that factor 1 captures the most generalized dimension of attitudes toward surveys, this factor advances into the multivariate modeling.

The other variables entered into the zero-order modeling are SES (including the dummy for no information), age, refusal on city directory, sample (1982 vs. 1985), and number of contacts reported on the response history. SES and age merit inclusion since these emerged from Table 1 as the two most serious sociodemographic biases when comparing respondents (given contact) to nonrespondents in the Waterloo surveys. The importance of refusal on city directory, as well as number of contacts on the response history, has been discussed already. Sample is included as a check on how accurately the response histories reproduce across the two samples.

The zero-order coefficients for the nonattitudinal predictors reveal,

---

[7] There is a marginally significant tendency for those interviewed by telephone (i.e., the 1985 Waterloo sample) to hold a higher tendency to have answered the last request for a telephone interview. That relationship is intuitively so sensible it is tempting to believe that it would reproduce in other data sets.

Table 4. Response History Analysis, Waterloo Survey Respondents

Predictors	Logistic Regression Coefficients (Dependent Variable = Proportion with No Refusals)				
		Multivariate			
	Zero-order	*1*	*2*	*3*	*4*
Intercept	—[a]	1.773	1.612	1.682	2.430**
Factor 1	.566**		.567**	.494**	.378
Factor 2	.368*				
Factor 3	†				
Factor 4	.236				
SES	−.012	−.012	−.011	−.010	†
No data, SES	†	.522	†	.422	.326
Age	−.019	−.023*	−.022	−.024*	−.023*
City directory refuser	−.916*	−1.484**		−1.316**	−1.205*
Sample	†				
Number of contacts	−.812***				−.714***

[a] See notes, Table 1.

Factor scores: by regression method, from Table 3/Sample: 1982 sample = 1, 1985 sample = 0/Contacts: summed from question sequence described in Table 3, range is 1 to 3/other predictors, see notes to Table 1.

N = 156 (the number reporting at least one prior contact).

first, virtually no ($p > .1$) effect of SES on response history. The effect of age takes the expected (from modeling in Table 1) direction that older people have higher probability of refusal, but fails significance as a zero-order effect. Refusers on the city directory carry high, but only marginally significant, probability of having refused at least one survey request. The next zero-order coefficient confirms that the response histories reproduce accurately across samples, in the sense that no variation in probability of having refused a survey is detectable across the two Waterloo samples. Finally, the zero-order coefficients underline the obvious reality that the greater the number of times a person is asked to participate in a survey the greater the odds that they will have refused at least one request.

The first multivariate model disentangles the effects of city directory refusal and noninformation on occupation. The two correlate, because the city directory by definition can provide no information on a refuser. A suppression is evident, with the marginally significant city directory refuser zero-order coefficient of .916 climbing to a .05 significant level of 1.484. Age now attains marginal (.1) significance, and maintains a marginal probability in most subsequent models.

The second multivariate model tests the extent to which the effect of attitudes toward surveys on response history is independent of sociodemographic factors. The fact that the net effect coefficient for factor 1 is unchanged until the third decimal place (.567 vs. .566) over the corresponding zero-order effect demonstrates that attitude stands virtually orthogonal to the sociodemographic variables.

Next (third multivariate model), city directory refusal reenters the modeling, giving our test of whether attitudes toward surveys are less favorable among those defined from external evidence as survey resistors. This likelihood is confirmed in the test since, sharing variance, the coefficients for both factor 1 and city directory refusal decline (comparing multivariate model 3 with models 1 and 2).

Finally, number of requests for survey response joins the equation, in multivariate model 4. While the zero-order effect that repeated requests enhance the probability of at least one refusal is mathematically self-evident, the multivariate relationship is not. If repeated requests for survey cooperation generated increasingly favorable attitudes to surveying, then controlling for number of requests would enhance the net strength of attitude as a predictor of response history. (Number of requests would suppress, correlating positively with probability of refusal, and negatively with *un*favorable attitude, while unfavorable attitude correlates positively with probability of refusal). But as it turns out, the greater the number of requests for an interview, the greater the unfavorability of attitude (as a Pearsonian correlation, $r = -.22$, $p = .003$), and for this reason controlling for contacts erodes the importance of attitudes, in multivariate model 4. Factor 1 falls below .1 significance, although its positive sign remains reproducible across each sample. The meaning of this and other results receives attention below.

CONCLUSION

The article began from the fact that while the volume and quality of survey on surveys data are increasing, so are reservations about the validity of such information. The central doubts have been that (1) since some appreciable incidence of nonresponse seems unavoidable on any noncompulsory modern survey, respondents on a survey on surveys may be self-selected from that segment of the public most favorably disposed toward the survey method; (2) there may simply be no attitude toward surveying, in the enduring, psychologically deep-seated, reproducible sense.

Our design entailed working at the municipal level, since here external information on sample listings is readily accessible. An analysis of nonresponse bias on surveys on surveys collected in Waterloo, Ontario in 1982 and 1985 revealed a socioeconomic profile not unlike those uncovered in most other surveys (e.g., Hawkins, 1975; Fitzgerald and Fuller, 1982)—an overselection of upper socioeconomic status, hints of low response among such social marginals as the aged and ethnoreligious minorities, and evidence that any correlate of accessibility will prove important, unless response is redefined to comprise no more than the result of confirmed contact.

On the more pointed question, "Do surveys on surveys self-select from believers?" evidence is also at hand: the variable identifying refusal on the city directory—our external measure of orientation toward surveying—performs rather more strongly as a correlate of reported refusal on *other* surveys than in predicting refusal on the survey on surveys. A survey organized around the topic of "What do people think of surveys?" seems to appeal to those opposed to surveying, as well as to the "professional respondents" who thrive on being interviewed on any topic. Thus, it is the elasticity of survey response according to topic, a theme so often stressed in the survey methodology literature (e.g., Heberlein and Baumgartner, 1978:449), that saves surveys on surveys from terminal selection bias.

The analysis of attitudes toward surveying began with a factor analysis of 12 items (some in turn summarizing subitems) from the Waterloo surveys. Four factors were identified by the factor analysis program, with the first conveying the most general meaning when factor loadings were interpreted at face value. The replicative design of the two Waterloo surveys permitted a test of factor stability. Since the surveys were collected three years apart, passing events such as an inaccurate election polling result can be dismissed as common determinants of the factor structure for the two samples. The first factor in the replication correlates highly across samples, enhancing confidence in the reality of the general attitude dimension detected by the factoring. Reproducibility for the secondary factors was not so strong.

The attitude dimensions were then assessed in a construct validity sense, from the argument that if attitudes toward surveying are important, they should correlate with response behavior. The attitude factor variables passed this test, in the believable pattern that the more general the factor the stronger the association with response history. It was discovered in addition that attitudes are related to exposure to surveying, in that the more times a person reports requests for survey cooperation, the more unfavorable is his or her attitude toward the method. It will be useful to see if other researchers can reproduce this potentially important pattern. If the relationship does generalize beyond two small samples collected in a mid-sized Canadian town, surveyors will wish to exercise renewed vigilance against oversurveying and against bogus surveys.[8]

The response history admittedly amounts to the least secure component of the design, since the history itself requires a validity check. Judging from the evidence presented above (note 6), response histories are easily reproducible across alternate question designs, but that does not eliminate the chance of recall errors, conscious and unconscious, in any such question.

---

[8] The depressant effect of contacts on response is not solely due to bogus surveying. The Waterloo respondents were asked: "Has a salesman ever posed as a surveyor in order to try to sell you something?" Including that variable in the computation by no means erases the relationship described in Table 4, model 4.

The findings (Table 4) of a noncorrelation between SES and prior refusal history illustrates the dilemma. Presumably the expected finding fails because the SES effect has already been exhausted by SES-biased selection into the sample of those consenting to participate in the Waterloo surveys. The more worrying possibility, however, is that the response history is so error-prone that a correlation between SES and response history is obscured. Given the low correlation between attitudes toward surveys and SES, the former interpretation seems the more plausible, but it would be satisfying to achieve external confirmation of the accuracy of the response history reports.

Taken as a whole, we believe the evidence presented here belies deeply pessimistic views about surveys on surveys. Indeed, evidence apart, the one part of the pessimistic account negates the other If only surface opinions about surveying exist in the public mind one can hardly speak of professional respondents, and vice versa. Read with due caution and with maximum recourse to external evidence to surmount the epistemological contradiction, surveys on surveys offer potentialities for the market researcher interested in industry image as well as to the academic interested in surveys as social phenomena.

References

Allen, Irving, and J. David Colfax
1968 "Respondents' attitudes toward legitimate surveys in four cities." Journal of Marketing Research 5:431–33.
Arnold, Rome G.
1964 "The interview in jeopardy: a problem in public relations." Public Opinion Quarterly 28:119–23.
Baxter, Richard
1964 "An inquiry into the misuse of the survey technique by sales solicitors." Public Opinion Quarterly 28:124–34.
Biel, Alexander L.
1967 "Abuses of survey research techniques: the phony interview." Public Opinion Quarterly 31:298.
Blishen, B. R., and H. A. McRoberts
1975 "A revised socioeconomic index for occupations in Canada." Canadian Review of Sociology and Anthropology 12:71–79.
Day, George S.
1975 "The threats to marketing research." Journal of Marketing Research 12:462–67.
DeMaio, T. J.
1980 "Refusals: who, where, and why?" Public Opinion Quarterly 44:223–33.
Fitzgerald, Robert, and Linda Fuller
1982 "I hear you knocking but you can't come in: the effects of reluctant respondents and refusers on sample survey estimates." Sociological Methods and Research 11:3–32.

Goudy, Willis J.
1977 Nonresponse Effects: Studies of the Failure of Potential Respondents to Reply to Survey Instruments. Monticello: Council of Planning Librarians.
Goyder, John
1985a "Surveys on surveys: limitations and potentialities." Paper presented at the 40th Annual AAPOR Conference, McAfee, New Jersey.
1985b "Polls on polls show mixed views on polls." Kitchener-Waterloo Record, May 17:A7.
Groves, R. M.
1979 "Actors and questions in telephone and personal interview surveys." Public Opinion Quarterly 43:190–205.
Groves, R. M., and R. L. Kahn
1979 Surveys by Telephone: A National Comparison with Personal Interviews. New York: Academic Press.
Hawkins, Darnell F.
1975 "Estimation of nonresponse bias." Sociological Methods and Research 3:461–88.
Hanushek, E., and J. Jackson
1977 Statistical Methods for the Social Sciences. New York: Academic Press.
Harman, Harry H.
1976 Modern Factor Analysis, 3rd ed. Chicago: University of Chicago Press.
Heberlein, Thomas A., and Robert Baumgartner
1978 "Factors affecting response rates to mailed questionnaires: a quantitative analysis of the published literature." American Sociological Review 43:447–62.
Kohut, Andrew
1986 "Rating the polls: the views of media elites and the general public." Public Opinion Quarterly 50:1–9.
National Research Council
1979 Privacy and Confidentiality as Factors in Survey Response. Washington, D.C.: National Academy of Science.
New York Times
1985 "Pollsters cite surveys indicating confidence in their work." May 20:B7.
Rogers, T. F.
1976 "Interviews by telephone and in person: quality of responses and field performance." Public Opinion Quarterly 40:51–65.
Rokeach, Milton
1968 "Attitudes." Pp. 449–57 in David L. Sills (ed.), International Encyclopedia of the Social Sciences. Volume I. New York: Macmillan and Free Press.
Roper, Burns W.
1986 "Evaluating polls with poll data." Public Opinion Quarterly 50:10–16.
Roper Organization
1985 Roper Reports, March/April.
Rugg, W. Donald
1972 "Interviewer opinion on the 'salesman as interviewer' problem." Public Opinion Quarterly 35:625–26.
Schleifer, Stephen
1985 "Trends in attitudes toward and participation in survey research." Public Opinion Quarterly 50:17–26.
Schwartz, Alvin
1964 "Interviewing and the public." Public Opinion Quarterly 28:135–42.
Sharp, Laure M., and Joanne Frankel
1983 "Respondent burden: a test of some common assumptions." Public Opinion Quarterly 47:36–53.
Singer, Eleanor
1984 "Public reactions to some ethical issues of social research: attitudes and behavior." Journal of Consumer Research 11:501–09.
Sjoberg, G.
1955 "A questionnaire on questionnaires." Public Opinion Quarterly 18:423–27.

Gender Differences in Policy Preferences: A Summary of Trends from the 1960s to the 1980s

ROBERT Y. SHAPIRO AND HARPREET MAHAJAN

THE "GENDER GAP" in the public's support for the first Reagan administration made political analysts more attentive than in the past to sex differences in political attitudes and behavior and to the long-term impact of the women's movement. Since the early days of survey research such differences were (or were thought to be) small and inconsequential. Men and women tended to exhibit similar aggregate patterns and trends in voting and political partisanship, although women tended to be less politically active and interested in politics.

Abstract Using 267 repeated policy questions (962 time points), we examine gender differences in policy choices and how they have changed from the 1960s to the 1980s. The average gender difference in preferences toward policies involving the use of force have consistently been moderately large. Sex differences in opinion toward other policies—regulation and public protection, "compassion" issues, traditional values—have been approximately half as large but they also warrant more attention than in the past. Our analysis suggests that the salience of issues has increased greatly for women, and as a result differences in preferences have increased in ways consistent with the interests of women and the intentions of the women's movement.

Robert Y. Shapiro is Assistant Professor of Political Science at Columbia University and a Research Associate at NORC, University of Chicago. Harpreet Mahajan is a Ph.D. candidate in Political Science at Columbia University. The authors owe special thanks to Kurt Veith for his collaboration in some of the research reported here. They are also grateful to Ethel Klein, Eleanor Singer, Ester Fuchs, Eric Smith, Tom Smith, Chaim Eyal, Benjamin Page, Garth Taylor, and the anonymous reviewers for their helpful comments and suggestions. The survey data and tabulations were obtained from the files of the Policy Change and Public Opinion Change project at NORC, which was originally supported by National Science Foundation Grant No. SES-7912969. Support for the research reported here was provided by the Columbia Council for Research in the Social Sciences; the responsibility for analysis and interpretation is the authors'. An earlier version of this article was presented at the 1984 Annual Meeting of the Midwest Association for Public Opinion Research, Chicago, Illinois.

Public Opinion Quarterly Vol. 50:42–61 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-42/$2.50

Significant changes, however, occurred. Not only were women less supportive than men of President Reagan but they had become less likely to consider themselves Republicans (e.g., Clymer, 1983; Schneider, 1984; Frankovic, 1982), and they had become more active politically than in the past and voted in greater numbers (see Wolfinger and Rosenstone, 1980; Verba and Nie, 1972; Welch, 1977; Klein, 1984; Andersen, 1975; Rapoport, 1981). Despite premature reports in November 1984 of a decrease in the gender gap since 1980, the evidence indicates that the sex difference in President Reagan's support was about the same in both elections and that the gender gap may have affected the outcomes of races for other offices (see Opinion Roundup, 1985:28–29; Klein, 1985; Keeter, 1985).

Less attention has been paid to gender differences in *policy preferences* than to voting. Public opinion surveys usually revealed no appreciable differences in choices for government policies, and these seemingly inconsequential findings were not extensively reported and summarized. They were mentioned, however, in voting and public opinion studies (Erikson et al., 1980:186; Monroe, 1975:96; Pomper, 1975; Campbell et al., 1960; Berelson et al., 1954), and breakdowns of responses by sex were often (and are still) routinely reported with survey results and in special compilations, such as those prepared for *Public Opinion Quarterly* (e.g., "The Polls" and the earlier "Quarterly Polls"; see also Cantril with Strunk, 1951).

Not surprisingly, there has been increasing interest in these differences in policy preferences because of their potential electoral importance.[1] Moreover, although some policy issues may be electorally more important than others, any changes that occur in the aggregate distributions of public preferences may ultimately influence the political agenda and the direction of public policy (see Monroe, 1979; Page and Shapiro, 1983). These opinions, to be sure, are not autonomous but may be shaped by events and objective conditions, and by processes of education, leadership, or manipulation (cf. Page and Shapiro, 1982, 1984). Gender and other differences may be important because some groups may be more politically powerful than others or their preferences may be the driving force behind changes in *mass* opinion. Gender differences may take on additional significance because of women's rising levels of political participation (and women simply constituting a large pool of eligible voters) and because of the greater attention they may pay to particular issues and government policies. As a result, even small sex differences in policy choices may no longer be inconsequential, especially if they occur across a wide array of issues around which coalitions can form.

---

[1] For example, Frankovic (1982) showed that women's lower level of support for Reagan could be explained by their greater opposition than men to belligerent positions on war and peace issues, arms control, and the uses of force in foreign policy. This assumes, of course, that these issues and any others were equally important influences on voting decisions for men and women. But this need not have been the case. Klein (1984) showed that some policy preferences for which there were no significant sex differences had different effects on candidate choices. Most notably, men and women had similar opinions toward the ERA and other policies pertaining to women's rights, but over time women appeared to be more likely to take these issues into account in voting. Thus the political importance of policy choices may go beyond the differences and similarities in the positions held by different people. On the other hand, such differences are most likely to occur on salient and important issues, and we would expect larger differences to produce larger effects. Moreover, it is also much more difficult for candidates and policy makers to determine the relative importance of different issues for different groups, *everything else being equal* (i.e., positions on other issues, partisanship; see Klein, 1984; Shabad and Andersen, 1979), than it is to judge simply the extent of differences in preferences.

What do we know about these gender differences? What differences have there been and what changes, if any, have occurred? We will briefly summarize the findings of past studies and then build upon them with an analysis of a larger set of additional data.

Previous Studies

Until recently there were no large collections and analyses of gender differences in various types of political attitudes and opinions. Previous studies examined such differences using a small number of surveys and opinion items, but their results did not lead to much other research or the need to synthesize a broad range of findings. The most thorough analyses were done by Smith (1984), the editors of *Public Opinion* magazine (Opinion Roundup, 1982), Schneider (1984), and Erskine (1971).

Smith (1984) reported that there have been rather striking and persistent gender differences in perceptions and opinions on matters concerned with force and violence, both on the domestic front and in our country's dealings with the rest of the world. Although he did not exclusively examine policy preferences, the differences that he found were compelling: men were more likely than women to support violent options in 87 percent of his readings, and the average gender difference was more than 9 percentage points, with individual differences ranging up to 30 percent.

The *Public Opinion* editors help fill out the picture about other issues by summarizing a sizable but undisclosed number of trends and patterns of differences since 1948. Their findings on violence are consistent with Smith's, except that the gender differences had increased since the mid 1970s. Moreover, they observed that differences had emerged or increased on "compassion" issues concerned with policies which aid the poor, the unemployed, the sick, and others in need, or which protect the well-being of people in general. Women were more supportive than men of such policies (cf. Cook, 1979).

Various authors have explained these emergent differences both by the way women have been socialized as caretakers and nurturers (e.g., see Gilligan, 1982; Stoper and Johnson, 1977; Chodorow, 1974) and by the rise of a women's political consciousness. Women may have extended the scope of their more traditional roles to encompass the problems of society as a whole (see Smeal, 1984; Klein, 1984; Friedan and Decter, 1982). The issues of concern to women thus might include policies to help those in need and also a somewhat distinct set of regulatory policies (e.g., environmental regulation, consumer protection, other safety issues) which are intended to protect the public. The increase in support for these policies would be more likely to occur among newer cohorts of women who, in contrast to earlier generations, came of political age after the emergence of the women's movement. Of course, some women may support the same policies not because of some new political consciousness but because of psychological and developmental differences between the sexes, which lead them to respond in certain ways to pressing new issues. That is, the new issues themselves may be the driving force, but these motivations are difficult to untangle.

Some of these same explanations can be applied to gender differences on force and violence issues and to opinions about policies that preserve and protect traditional values, the home, the family, children, religion, the neighborhood or immediate community, and the order and stability of everyday life. Deviant behavior and personal excesses were to be discouraged and punished. This support for traditional values, broadly defined, is a form of social conservatism which can be compared and contrasted with women's economic liberalism on compassion issues.[2]

---

[2] A single, broadly defined liberal–conservative ideological continuum almost certainly cannot be applied here to women, if it ever could (cf. Nie et al. 1979; Miller and Levitan, 1976; Converse, 1964). It is possible that men and women—or different generations of them—organize their attitudes and opinions in different ways. Compassion, traditional values, regulation and protection, and force and violence may be the major dimensions which help structure women's attitudes. Others may apply to men. The evidence thus far about this type of gender difference (using factor analysis) has been limited and inconclusive; it is uncertain whether men and women systematically organize their political views and policy choices in the same way, if at all (see Goertzel, 1983). What we think is likely is that the dimensions described above do not all apply to all women, but that different ones are important to different women. Thus socialization processes and the consciousness associated with the women's movement can have quite varied influences on political attitudes and opinions.

Some women may perceive women's rights issues themselves, and especially abortion, as a challenge to traditional values. For other women these issues may be distinct, creating only minor tensions with prior values. According to the *Public Opinion* editors, gender differences on traditional values issues appeared to be diminishing, and on women's rights issues there were generally no differences and no changes over time.[3]

Erskine (1971) examined sex differences in attitudes toward women's roles in politics and society (all not policy issues, strictly speaking) for the period from the mid 1930s to the 1970s. She found that prior to the 1960s women were much more supportive than men of expanding women's participation in economic and political life (e.g., employment opportunities, holding public office), although by the 1960s such differences had diminished or even reversed (Erskine, 1971:275).

Since so few studies have been done and since only Smith and Erskine provided statistical estimates of gender differences, many questions remain. We will examine the magnitudes and directions of gender differences for the types of policy issues which we have described, and will appraise the changes that occurred from the 1960s to the early 1980s. By this period women had shifted away from the Republican Party and remained proportionately less Republican than men (Schneider, 1984:1242). These were also the years in which the women's movement first had an opportunity to influence political attitudes and behavior.

Data and Methods

To examine gender differences in preferences we will compare responses to policy questions which have been asked in national surveys. The opinion items that we report here are from surveys conducted by NORC (NORC General Social Surveys (GSS), 1972 through 1983; Davis, 1982 and 1983); Gallup (the 1964–1983 volumes of the *Gallup Opinion Index/Report*); Louis Harris and Associates (from press releases and the *Harris Survey Yearbooks* for 1970, 1971, 1972, and 1973); the Opinion Research Corporation, surveys done for the Advisory Commission on Intergovernmental Relations from 1973 to 1982 (1975–1982), and the Survey Research Center/Center for Political Studies, including the 1952–1978 American National Election Studies (Miller et al., 1980).

---

[3] Cf. Schneider (1984:1243), who states that women were less supportive than men of women's rights 10 years ago, but that the difference has disappeared.

The tabulations from these surveys are to our knowledge the largest set of separate opinion trends for men and women yet assembled. The data consist of 267 different repeated policy questions, and they cover a broad range of foreign and domestic policies. Approximately half contain three or more time points, and they total 962 time points in all for examining gender differences. The data are not a sample, strictly speaking, but they deal with a wide variety of laws, different types of government spending, regulations, acts of government officials, and outcomes, although there are disproportionately more domestic (210 items, 780 time points) than foreign affairs questions (57 items, 182 time points).[4]

One substantive and methodological issue warranted very special attention: how to treat "don't know," "no opinion," and "not sure" responses? We did our analysis both including and excluding them, and we found that the different treatments of don't knows did not affect our results. Past research and some compelling theoretical considerations, however, suggested that the don't knows were important in their own right. Women have always had higher proportions of these responses than men, but there is evidence that this is much less so for new generations of women (see Smith, 1984; Rapoport, 1982). In addition, Shapiro (1982) reported that the proportion of don't knows in general have decreased over time. This conceivably can be attributed in part to the drop in survey response rates, perhaps of apathetic and unattentive members of the public, or to the public getting used to surveys and feeling obligated to provide answers. But the don't knows are also a crude measure of the salience and importance of political issues (cf. Page and Shapiro, 1983; Smith, 1981). The decline in these responses may indicate a rise in the visibility and significance (both personal and public) of policy issues, especially new ones such as the social issues which emerged in the 1960s, and this may have been facilitated by rising levels of education and increases in the amount of information that was easily accessible to the public through television and other media (see Nie et al., 1979). Changes in policy preferences among women may be related to the number of don't know responses, which are affected by attentiveness to politics and interest in issues. Consequently, we estimated gender differences in don't knows and how they changed since the 1960s.

---

[4] The data provide a way of examining the reliability and robustness of the other authors' findings. Some of the items which they used are contained in our data set (e.g., perhaps as many as half of our 164 cases dealing with force and violence were analyzed by Smith (1984)), but there are probably more than 700 additional time points as well.

Analysis and Findings

FORCE AND VIOLENCE

The first question we examined was how different were issues involving the use of force or violence from all other issues. We classified each survey question according to whether it (1) clearly dealt with force or violence, (2) clearly did not, or (3) was ambiguous. For example, issues such as defense, troop levels abroad, capital punishment, and gun control fell into the first category, whereas the space program, the country's activeness in international affairs, the courts' treatment of criminals, and abortion were in the third. As would be expected, the largest number of survey items were in the second category.[5]

Within each category, we calculated the average gender difference (percentage point difference), and we also determined the proportion of time points for which individual gender differences (each time point for a time series) were a substantively and statistically significant (at better than the .05 level) 6 percentage points or more.[6] In one opinion series on abortion, for example, we found 10 successive readings during the period from 1972 to 1983 in the NORC GSS (there were no 1979

---

[5] We were aware that here and later in our analysis there would be some temptation to reclassify issues based on afterthoughts and patterns of findings, especially for policies which tapped several important dimensions at the same time. Doing this leads to larger gender differences for certain categories of issues and strengthens our findings and conclusions. We were uncomfortable, however, with altering our original groupings of issues after-the-fact, since our initial judgments had at least equal substantive merit. Moreover, if the data supported certain theoretical arguments, then the findings should be robust enough to weather the difficulties encountered in categorizing a small number of issues. We leave any reinterpretations of our data to the reader.

[6] To determine the size of each gender difference, we dichotomized the response categories to find the largest difference, making sure, however, in questions with three or more ordinal categories that the *balance* of opinion was different (i.e., that opinion was not merely more polarized for women than men, or the reverse). For the SRC/CPS items which we took from Miller et al. (1980), we divided the reported Percentage Differences Indexes (PDI's) by 2 to make the differences in PDI's roughly comparable to differences in relative frequencies. Don't Know, No Opinion, and Not Sure responses were excluded in the data reported here, except for the SRC/CPS questions.

We initially thought that summing the differences of the aggregate responses (allowing positive and negative ones within a time series to offset each other), taking the absolute values of these sums, and then averaging these absolute values (dividing by the total number of time points involved), would produce small numbers that would be no more than sampling errors for cases in which there were no differences. That is, we expected minuscule average differences because we thought there would be a lot of small and randomly fluctuating positive and negative differences which would simply cancel each other out and produce smaller and essentially meaningless averages. Although the final sums and averages were relatively small, the individual gender differences within each opinion series did not vary much over time and tended *not* to switch signs. Since these differences persisted, it was clear that they did not merely result from sampling error. Thus, allowing for positive and negative gender differences within a given question produced averages that were meaningful and revealed real gender differences.

Table 1. Gender Differences on Policy Issues

	1A. Average Percentage Point Differences				*1B. Percentage of Differences of at Least 6 Percentage Points*			
	Survey Organization				*Survey Organization*			
Issues	*NORC-GSS*	*Gallup*	*Other*[a]	*Total*	*NORC-GSS*	*Gallup*	*Other*	*Total*
Force/ violence	7.6 (52)	8.9 (70)	7.3 (42)	8.1 (164)	71% (52)	66% (70)	67% (42)	68% (164)
Ambiguous	4.4 (112)	5.1 (93)	3.7 (39)	4.6 (244)	27 (112)	40 (93)	33 (39)	33 (244)
Nonforce/ violence	3.8 (215)	4.5 (166)	3.6 (173)	4.0 (554)	33 (215)	38 (166)	24 (173)	31 (554)
Total	4.5 (379)	5.6 (329)	4.3 (254)	4.8 (962)	36 (379)	44 (329)	32 (254)	38 (962)

[a] Includes HARRIS, SRC/CPS, and ACIR/ORC Surveys (80, 123, & 51 time points, respectively). Don't Know, No Opinion, and Not Sure responses have been excluded here and in all other tables for all surveys, except the SRC/CPS surveys.

and 1981 surveys) which showed that men were more supportive than women (by 5, 10, 2, 1, 6, 4, 3, 2, 4, and 7 percentage points) of abortions for women who did not want to have more children. The 4 percent average difference for these represents a true and persistent difference in a certain direction that may have been as great as 10 percent. This type of data analysis admittedly only provides rough estimates, particularly when we take averages over different policies and survey questions. But, as we will see, it can be quite revealing.

Our first results are shown in Table 1, which reports the findings by survey organization. Table 1A presents the average differences for those issues which involve force, those that do not, and those which are ambiguous. Table 1B lists the percentage of time points in which there were gender differences of at least 6 percentage points. The findings are similar for the different survey "houses," which follow different procedures and ask different questions (see Smith, 1978). The gender differences for force issues averaged 8 percentage points, with two-thirds of them over 6 percent, and as we shall see, this almost completely reflects men selecting more violent options.[7] These differences are twice as large, and the 6 percent changes twice as frequent, as those for nonforce issues. The ambiguous policies fall in between, but they are clearly much more like the nonforce issues, suggesting that rather sharp distinctions on the force–violence dimension may be necessary to produce gender differences.

Given the common wisdom about there having been essentially no sex differences in attitudes, the differences that we found for nonforce issues may be somewhat surprising. Six percent differences occurred in one of every three cases, and the average difference of more than 4

---

[7] Some of the exceptions, which may have reflected genuine sampling errors, were small differences on military matters and on permitting a person to hit another person under certain conditions.

Table 2. Average Percentage Point Differences in Specific Policy Directions (% Women − % Men)

	Survey Organization							
Issues	*NORC-GSS*		*Gallup*		*Other*		*Total*	
Opposition to force-domestic	7.6	(39)	12.3	(26)	7.2	(15)	9.0	(80)
Opposition to force-foreign	6.1	(13)	5.6	(44)	7.1	(27)	6.2	(84)
Support for compassion	3.6	(65)	2.0	(40)	3.8	(68)	3.3	(173)
Support for regulation, protection	1.0	(9)	10.0	(21)	2.8	(12)	5.8	(42)
Support for traditional values	3.0	(133)	1.8	(27)	0.0	(39)	2.3	(199)
Support for women's equality, abortion	−1.6	(76)	−3.3	(55)	−1.9	(24)	−2.2	(155)
Other issues (absolute magnitude)	7.9	(44)	3.9	(116)	3.7	(69)	4.6	(229)
Total (absolute magnitude)	3.9	(379)	4.7	(329)	3.5	(254)		

percentage points may not be trivial, if it reflects consistent directions of gender differences and sizable differences for some issues.

The directions of these differences are captured in Table 2. The larger differences for the Gallup domestic-force issues can be attributed to gun control and capital punishment trends for which there are substantial differences.[8] The differences are smaller on foreign and defense matters, although these include few questions on the "gender gap" issues of missile systems and relations with the Soviets. In addition, questions have been included on expanding civil defense (e.g., bomb shelters), which women support more than men by 10 percent or more, despite the possible argument that such defenses might encourage the use of force. Our best judgment is that the gender differences are probably about the same for foreign versus domestic violence, with differences for issues such as gun control, capital punishment, defense spending, the withdrawal of troops from Vietnam, mining the harbors of North Vietnam, and providing arms to the Israelis and the Arabs as high as 15 to 20 percent.[9]

---

[8] E.g., in response to a 1981 Gallup question ("In general do you feel that the laws covering the sale of handguns should be made more strict, less strict, or kept as they are now?") 75 percent of the women respondents versus 56 percent of the men favored stricter laws. Also in 1981, 51 percent of the men versus 38 percent of the women favored capital punishment for treason ("Are you in favor of the death penalty for persons convicted of treason?"). See Smith (1984).

COMPASSION ISSUES

The compassion issues are largely about jobs, income redistribution, and other economic policies to help the poor and targeted group. These include spending on social welfare, education, health, and programs to assist blacks and poor states and central cities, and inflation and unemployment policies.[10] The average gender difference for these policies is just over 3 percent, less than half that of the force issues. The difference is less for the Gallup cases (lowering the overall average), and the issues responsible for this deal with balancing the federal budget: women tended to support a balanced budget by as much as 8 to 10 percent more than do men, despite the possible implications of this for programs that benefit the poor, the elderly, and the like.[11] But women were more supportive of a guaranteed annual income, wage-price controls, equalizing wealth, guaranteeing jobs, government-provided health care, student loans, and rationing to deal with scarce goods.[12]

REGULATION AND PROTECTION

There are larger differences for policies which regulate and protect consumers, citizens, and the environment. The NORC-GSS revealed no difference for spending on the environment, although here the overall average may conceal a change which occurred: by the 1980s perhaps 5 percent more women than men supported spending.[13] In addition, the SRC/CPS and ACIR/ORC surveys include items on the government's general power to protect citizens, perhaps bringing to mind images of "big government" which should be opposed while individual safety regulations are supported. For example, there is overall no appreciable difference in responses to a question about whether the government has too much power,[14] but there are striking differences for particular types of regulations and safeguards. Twenty percent more women than men have expressed opposition to nuclear power plants, and women have also shown more support for the 55 mile per hour speed limit, fines for people who do not wear seat belts, jail terms for drunk drivers, and banning cigarette advertising and sales.[15]

---

[9] E.g., September 1973 Gallup: "There is much discussion as to the amount of money the government in Washington should spend for national defense and military purposes. How do you feel about this: Do you think we are spending too little, too much, or about the right amount?" Percent "too much": 58% women, 45% men. June 1969 Gallup: "Some people think that the time has come to begin to reduce month by month the number of U.S. soldiers in Vietnam. How do you feel—do you think the time has come to do this, or not? Percent "no": 38% men, 21% women. See Smith (1984).

[10] E.g., 1976 NORC-GSS: "We are faced with many problems . . . are we spending too much, too little, or about the right amount on . . .?" Solving the problems of the big cities—percent "too little": 45% men, 51% women. Improving the nation's education system—percent "too little": 49% men, 54% women. Improving the conditions of blacks—percent "too much": 32% men, 23% women. Welfare—percent "too much": 66% men, 60% women.

Taxes were not included in this category of issues; women consistently showed more opposition than men to federal income taxes.

[11] Of course, one may speculate that some women associate this issue with good management or care for the family or household budget.

[12] E.g., January 1969 Gallup: "As you may know, there is talk about giving every family an income of at least $3200 a year, which would be the amount for a family of four. If the family earns less than this, the government would make up the difference. Would you favor or oppose such a plan?" Percent "favor": 31% men, 38% women. September 1980 Gallup: "Would you favor or oppose having the government bring back wage and price controls?" Percent "favor": 52% men, 61% women.

The traditional values issues include pornography, the use of drugs such as marijuana, prohibition, prayer in schools, sex education, providing birth control for teens, and homosexuality and other alleged forms of deviant behavior. Women have consistently tended to uphold conservative values but the overall average difference is small.[16] The average for the NORC-GSS is highest because of the inclusion of several of Stouffer's tolerance/civil liberties items, in response to which women still express more support than men for limiting freedom of speech for would-be threats to society such as Communists, socialists, or atheists (see Stouffer, 1967; Nunn et al., 1978; McClosky and Brill, 1983).[17] We also included civil rights issues in this category, because these questions usually dealt with intermarriage, housing, and schools.

---

[13] See footnote 10 for question wording: " . . Improving and protecting the environment." In 1974, 5 percent more men than women responded (spending) "too little," whereas in 1978, 6 percent more women than men said "too little," 5 percent in 1980, and 2 percent in 1982.

[14] E.g., May 1978 ACIR/ORC: "Which of these statements comes closest to your view about government power today?" Percent "Federal government has too much power": 40% men, 41% women.

[15] E.g., February 1981 Gallup: "Do you favor or oppose keeping the present 55-mile-per-hour speed limit on the highways of the nation." Percent "favor": 72% men, 81% women. June 1981 Gallup: "Do you think there should or should not be a complete ban on cigarette advertising?" Percent "should be": 39% men, 53% women.

[16] E.g., March 1970 Gallup: "Do you think they (birth control pills) should be made available to teenage girls?" Percent "no": 72% men, 87% women. June 1980 Gallup: "Do you think the use of marijuana should be made legal, or not?" Percent "yes": 33% men, 20% women.

[17] E.g., 1982 NORC-GSS: "Now I would like to ask you some questions about a man who admitted he is a communist. Suppose this admitted communist wanted to make a speech in your community should he be allowed to speak, or not?" Percent "yes, allowed to speak": 62% men, 55% women.

Here differences varied depending on the issue. Overall there is no consistent and clear pattern, although if one takes into account compassion issues and positions concerning the general social and economic welfare of minorities, women perhaps tend to be slightly more liberal than men. When the civil rights issues are discounted among the traditional values policies, the average gender difference roughly doubles in the conservative direction.

These values may clash with abortion and with various facets of women's rights and equality. As we see in Table 2, women are overall more conservative on women's rights issues, which is almost completely due to men being consistently more pro-choice on abortion (based on questions concerning support for abortions under varying circumstances, such as rape, potential birth defects, danger to the woman's health). On other women's issues there are few clear and consistent differences.[18]

For the remaining (unclassified) issues Table 2 reports the average absolute magnitude of the differences. The average is 4.5 percent, and there are some large and noteworthy differences, especially among the NORC-GSS questions. For instance, support for wiretapping was more than 10 percentage points higher for men than women (women perhaps seeing this action as a forceful intrusion upon people's homes and lives), and women were more opposed (by over 15 percent) to spending on the space program (perhaps not so much because of its potential role in national defense as because spending on it used resources that might be used for other purposes). Some issues may tap several policy or ideological dimensions that are important to women, but it may be difficult to determine, a priori, their relative importance.[19]

Trends in Gender Differences

Thus far we have confirmed that there are gender differences on the force dimension and other significant variations as well. Women's concerns are a mixture of economic liberalism, government activism, and traditional social conservatism (with some social liberalism as well).

---

[18] This absence of a gender difference, however, obscures the possibly divergent sources of these preferences. The positions that men take on women's issues may be more likely to involve symbolic politics and liberalism, whereas women's choices may be more closely related to their experiences (Klein, 1984). As suggested earlier, men and women may be structuring their attitudes toward these issues in different ways.

[19] Other issues on which gender differences occur are campaign and election reforms, including the lengths of terms of office (with women more in favor of reforms), and policies toward labor and union activities. These differences cannot as readily be accounted for as the others.

Table 3. Changes in Gender Differences for Different Types of Issues (% Women − % Men)

	Time Period							
Issues	*1964–71*		*1972–76*		*1977–83*		*Total*	
Force-domestic	12.2	(14)	8.3	(29)	8.4	(37)	9.0	(80)
Force-foreign	7.2	(23)	5.7	(27)	5.9	(34)	6.2	(84)
Compassion	3.2	(29)	2.2	(55)	4.1	(89)	3.3	(173)
Regulation, protection	3.0	(5)	4.9	(12)	7.2	(25)	5.8	(42)
Traditional values	−0.9	(38)	2.2	(84)	3.6	(77)	2.3	(199)
Women's issues	−3.6	(8)	−2.1	(51)	−2.3	(53)	−2.2	(155)
Other	5.1	(73)	4.5	(71)	4.8	(85)	4.6	(229)
Total (absolute magnitude)	4.6	(190)	3.6	(329)	4.4	(443)		

But did these differences simply persist from the 1960s to the 1980s? The rough trends are presented in Table 3, in which we compare the average differences for three time periods, 1965–71, 1972–76, and 1977–83. In general the changes are modest.[20] The gender differences for domestic-force issues show an average drop of 4 percent from the 1964–71 period to 1977–83. This occurred through a small narrowing of the differences on capital punishment, gun control, and urban unrest, which still remained moderately large. Women may have been responding to higher crime rates, especially those crimes of which women are more likely to be victims (and perhaps increases in gun ownership among women contributed to their shift in opinion against gun control). Foreign-force showed a slight drop (probably distorted by civil defense policies and the omission of the space program and recent foreign and defense issues). There was no change from 1972–76 to 1977–83, and our best judgment is that there has been no change in gender differences on force issues taken as a group.

The remaining issues, except for those unclassified, show more interesting changes. For the compassion issues and more so for regulation-protection we find increases in gender differences. The differences for the latter are as big as those for all the force issues. One explanation for this is that after 1976, fiscal problems and (threats of severe) budget cuts in social spending and proposals for deregulation, despite visible safety and environmental problems, began to have an impact on public opinion, especially during the Reagan administration; and women reacted to this differently than did men.[21] The results are consistent with arguments about women's societal interests, although the table shows that the trends for compassion and regulation issues are not paralleled by the war and peace and other force issues, as some authors might have expected (cf. Smeal, 1984). Nor do the shifts for the compassion and protection issues fully mirror each other. They moved differently from the 1964–71 to the 1972–76 periods, but then shifted in tandem during the subsequent 1977–83 period, which is the crucial one. What may have happened here and in the cases of other issues as well was that the movements in aggregate opinion for different issues were determined by shifts in opinion by *different* women. Any questions about the characteristics of these women need to be answered elsewhere at the individual level, although we can speculate that at the extremes are women committed to ideas associated with the women's movement versus those who staunchly defend more traditional attitudes.

---

[20] For a single trend, a change in subgroup differences of about 10 percentage points might be considered substantively and statistically significant. Shapiro and Page (1984) found that diverging or converging trends of this sort are infrequent and less likely to occur for sex than for other population characteristics. In Table 3, however, we are comparing trends for a *large number of issues*, and smaller but consistent differences for them can be politically important. Cf. Taylor (1980) concerning pooled trends.

This becomes more apparent when we examine the traditional values issues. In contrast to compassion and regulation, women were more conservative on these matters, and the gap appeared to widen. The increase in these differences, however, is largely attributable to the different mix of issues included in this category, especially in the first and last periods. Several Gallup and Harris civil rights trends are in the 1964–71 period, and the NORC-GSS (Stouffer) tolerance items and questions on drugs, alcohol, and pornography are in the later time periods. In contrast, there may have been a slight drop for women's issues, but women still appear to be more conservative due to differences in opinion toward abortion.

Thus, a large number of women came to be even more liberal than men on compassion and regulation issues, while they or other women moved in a conservative direction or did not change their preferences on certain traditional values or social issues. Table 3 summarizes these dynamics but it offers no evidence about who altered their preferences and how this change came about. Women's greater concern than men for certain needy groups and for the well-being of society may have occurred because of their upbringing and other experiences in life. For some women the policies supported by the women's movement may have contributed directly, or the movement's impact may have occurred indirectly as it made politics and issues more salient and important for women than in the past. This may have occurred without having a uniform effect on women's attitudes. Women who wanted traditional values protected, especially women in older cohorts, may have been led to form and express their preferences. Some may have reacted negatively to liberal positions taken by participants in the women's movement which they perceived as extreme.

---

[21] As the public as a whole became more conservative during part of this period, this shift was greater for men than for women.

We would expect new generations of women to behave in a distinct way, or at least that the sources of their attitudes and behavior would be different from earlier cohorts. They are better educated and are more likely than members of past generations to be in the workforce (and for longer periods) and to experience more economic independence. They have been brought up in a society in which the women's movement has had a chance to exert direct or indirect influence. We would anticipate, then, that politics and issues would have become more important than in the past for women, in comparison to men, because of social and political influences on all women or as new cohorts became a larger part of the population. These competing explanations, of course, are not mutually exclusive; several processes can be at work at the same time.

Decline in Don't Know Responses

Rapoport (1982) examined some evidence concerning this. He observed that new cohorts of women gave fewer don't know responses to questions about politics asked in national surveys, and this very likely reflected the greater visibility of issues and politics for women and an increase in the attention that they paid to them. If the changes that we found can be attributed to the greater salience of these issues among women of all ages, or at least among younger cohorts (which we cannot untangle from our aggregated data), we would expect a decrease in the percentage of don't know responses on those issues for which there were changes in gender differences, that is, for compassion issues, regulation and protection, and traditional values.

We found that don't know responses have declined since the 1960s for both men and women (see Shapiro, 1982:80–81). Looking at the data by survey organization, since proportions of don't knows are highly sensitive to "house effects" (Smith, 1978), we find that the Gallup don't knows for men dropped from 8.1 percent ($n = 110$) to 7.3 percent ($n = 59$) and 7.4 percent ($n = 160$) when we compare the three time periods; for women the figures were 12.9, 11.1, and 9.9 percent. For men in the Harris surveys we found 12.2 percent ($n = 31$) for 1964–71 versus 9.2 percent ($n = 49$) for the later periods, and 18.3 to 12.5 percent for women. The ACIR/ORC survey showed 12.1 percent ($n =$

Table 4. Changes in Gender Differences in "Don't Know" Responses for Different Issues (% Women − % Men)

	Time Period							
Issues	*1964–71*		*1972–76*		*1977–83*		*Total*	
Force-domestic	1.7	(11)	1.9	(23)	2.0	(32)	1.9	(69)
Force foreign	5.2	(18)	5.3	(24)	5.7	(32)	5.4	(74)
Compassion	5.8	(17)	3.9	(47)	3.1	(69)	3.7	(133)
Regulation, protection	4.0	(1)	3.8	(10)	2.7	(23)	3.1	(34)
Traditional values	3.4	(24)	0.9	(73)	0.7	(75)	1.2	(172)
Women's issues	−1.3	(7)	−0.2	(45)	0.7	(92)	0.4	(144)
Other	6.8	(63)	4.5	(67)	3.8	(83)	4.9	(213)
Total	5.1	(141)	2.6	(289)	2.4	(409)	2.9	(839)

17) for men in the 1972–76 period versus 9.6 (n = 34) for 1977–83, and 18.4 versus 13.6 percent for women. The NORC-GSS showed smaller changes in don't knows.

Issues became more salient for everyone but especially for women, judging from the larger changes in don't knows. One possible reason for the smaller change among men is that they began with fewer don't knows to begin with and that there may be some kind of floor effect. That is, some of these responses may not occur because of the low salience of issues but rather because of the difficulty that politically attuned respondents may have in making their opinion fit one of the categories offered in closed-ended survey questions. These don't knows do not necessarily decline as issues become more important to people, and the don't knows among the men may have been of this type. Women may have only more recently neared the floor level of "educated" don't know (see Smith, 1981). Still, the importance of the decline in don't knows and the increase in the salience of issues among women is not diminished.

Table 4 summarizes the changes in the gender differences in don't knows. Historically there has been about a 3 percent difference, according to what we know from the survey research literature (Smith, 1984:383). That is approximately what we find (2.9 percent), but from the 1960s to the 1980s the difference has declined from 5 percent to approximately half that. The largest difference remains for foreign policy issues (force and several other issues). Not surprisingly, the smallest difference is for women's issues, men often having had more don't knows.

The overall drop in don't knows indicates that women have been catching up in terms of the attention that they pay to specific issues. It is possible that women were always as attentive as the data show for the most recent period, but that they were reluctant to express their views. They may have deferred to the opinions of men, or perhaps the process by which this occurred involved the hesitance (or intimidation) that people may feel if they think that their opinion is in a minority; "pluralistic ignorance" and the "spiral of silence" might have played a part (cf. Noelle-Neumann, 1974). The women's movement may have helped overcome this by legitimating and encouraging women to express their positions whether they were in the majority or not, and by the same token whether or not they supported the political positions of the movement itself.

Alternatively, the gender differences may have appeared to change because men reached the level at which don't know responses are no longer sensitive to changes in the salience of issues. In the 1960s and 1970s issues played a bigger part in *everyone's* lives. It was a dynamic period in American history. Issues such as civil rights, political protest, the Vietnam war, crime and other matters involving social control excited the nation. The media, especially television news, were there to cover the relevant events and public debate and to present their reports to a citizenry that was better educated than in the past (e.g., see Nie et al., 1979).

The findings in Table 4 and other evidence suggest that something more than the appearance of new and salient issues was affecting women. The don't knows declined for men and women on all issues, but the gender differences were more likely to decrease for issues on which the gender differences in preferences increased. Moreover, these were the issues for which opinions changed in directions that were consistent with the kinds of interests and concerns that women—perhaps very likely not the same women—may have already had or developed. For the compassion, regulation, and traditional values issues, the gender differences in don't knows were associated with *increases* in differences in preferences. In contrast, there was no change in the differences in don't knows for the force issues, while there was a decline or no change in the differences in policy choices.

If the floor effect argument provides a sufficient explanation, then the decrease in gender differences in don't knows should have occurred for all issues. It did not occur for the foreign-force issues, even though the large difference at the start suggests that these issues should have been less sensitive to floor effects. And if no policies became more important to women, the decline in don't knows should not have led to larger sex differences on particular issues. Thus, women's psychological involvement with questions of government policy and women's policy choices themselves may have been influenced distinctively by changes in society and politics.

Conclusion

Our analysis improves our understanding of gender differences in policy preferences and how they have changed from the 1960s to the present. During this period there were persistent differences in preferences. The average gender difference in preferences toward policies dealing with the use of force and violence have consistently been moderately large. Sex differences in opinion toward other policies have been approximately half the size, but they also merit more attention than they have been given in the past. The next largest differences occur for regulation and other protective policies and for compassion issues since the mid 1970s, which may be a consequence of social and political changes. The salience of these issues for women increased more than it did for men, and gender differences in opinions increased in ways suggested by the interests which women have had and consistent with the intentions of the women's movement. Similarly, women as a group also supported the protection of conservative traditional values more than did men. The impact of the women's movement is less clear here. On issues such as abortion and women's rights, women's greater conservatism probably diminished, but there are still small but persistent gender difference in opinions toward abortion.

As we get further into the 1980s and the second term of the Reagan administration, we will be able to find out whether the differences reported here remain or whether the trends continue or reverse. Because of ongoing changes in economic, social, and political life in the United States, such differences in policy preferences are likely to be looked at more closely in the future than they have been in the past. And it will be interesting to see how these differences affect electoral politics and the policy-making process.

References

Advisory Commission on Intergovernmental Relations
Changing Public Attitudes on Governments and Taxes. Washington, D.C.: ACIR. 1975, 1976, 1977, 1980, 1981, 1982 Vols.
Andersen, Kristi
1975 "Working women and political participation, 1952–1972." American Journal of Political Science 19:439–55.
Berelson, Bernard R., Paul F. Lazarsfeld, and William N. McPhee
1954 Voting. Chicago: University of Chicago Press.
Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes
1960 The American Voter. New York: Wiley.
Cantril, Hadley, with Mildred Strunk
1951 Public Opinion 1935–1946. Princeton, NJ: Princeton University Press.
Chodorow, Nancy
1974 "Family structure and family personality." In Michelle Zimbalist Rosaldo and Louis Lamphere (eds.), Women, Culture, and Society. Stanford, CA: Stanford University Press.

Clymer, Adam
1983 "If anything, gender gap is becoming even wider." The New York Times. Sunday, December 11, p.E5.
Converse, Philip E.
1964 "The nature of belief systems in mass publics." Pp. 206–61 in David E. Apter (ed.), Ideology and Discontent. New York: The Free Press.
Cook, Fay L.
1979 Who Should Be Helped? Beverly Hills, CA: Sage Publications.
Davis, James A.
1982 &
1983 General Social Surveys, 1972–82 and 1972–83: Cumulative Codebook. NORC, University of Chicago.
Erikson, Robert S., Norman R. Luttbeg, and Kent L. Tedin
1980 American Public Opinion: Its Origins, Content, and Impact. 2nd ed. New York: Wiley.
Erskine, Hazel
1971 "The polls: women's role." Public Opinion Quarterly 35:275–90.
Frankovic, Kathleen
1982 "Sex and politics—new alignments, old issues." PS 15:439–48.
Friedan, Betty, and Midge Dector
1982 "Are women different today?" Public Opinion April/May:20, 41.
Gilligan, Carol
1982 In a Different Voice. Cambridge, MA: Harvard University Press.
Goertzel, Ted George
1983 "The gender gap: sex, family income and political opinions in the early 1980s." Journal of Political and Military Sociology 11:209–22.
Harris, Louis, and Associates, Inc.
The Harris Survey Yearbook of Public Opinion, 1970–1973. 4 vols. New York: Louis Harris and Associates, Inc., 1971, 1975, 1976.
Keeter, Scott
1985 "Public opinion in 1984." Pp. 91–111 in Gerald Pomper et al., The Election of 1984: Reports and Interpretations. Chatham, NJ: Chatham House.
Klein, Ethel
1984 Gender Politics. Cambridge, MA: Harvard University Press.
1985 "The gender gap: different issues, different answers." The Brookings Review 3:33–37.
McClosky, Herbert, and Alida Brill
1983 Dimensions of Tolerance. New York: Russell Sage Foundation.
Miller, Warner E., and Teresa Levitin
1976 Leadership and Change: Presidential Elections from 1952–1976. Cambridge, MA: Winthrop Publishers.
Miller, Warner E., Arthur H. Miller, and Edward J. Schneider
1980 American National Election Studies Data Sourcebook, 1952–1978. Cambridge, MA: Harvard University Press.
Monroe, Alan D.
1975 Public Opinion in America. New York: Dodd, Mead.
1979 "Consistency between public preferences and national policy decisions." American Politics Quarterly 7:3–19.
Nie, Norman H., Sidney Verba, and John R. Petrocik
1979 The Changing American Voter. Enlarged ed. Cambridge, MA: Harvard University Press.
Noelle-Neuman, Elisabeth
1974 "The spiral of silence: a theory of public opinion." Journal of Communication 24:43–51.
Nunn, Clyde Z., Harry J. Crocket, Jr., and J. Allen Williams, Jr.
1978 Tolerance for Nonconformity. San Francisco: Jossey-Bass.
Page, Benjamin I., and Robert Y. Shapiro

1982 "Changes in Americans' policy preferences, 1935–1979." Public Opinion Quarterly 46:24–42.
1983 "Effects of public opinion on policy." American Political Science Review 77:175–90.
1984 "Presidents as opinion leaders: some new evidence." Policy Studies Journal 12:649–61.

Pomper, Gerald M.
1975 Voters' Choice: Varieties of American Electoral Behavior. New York: Dodd, Mead.

Public Opinion
1982 "Opinion roundup: women and men: is a realignment under way?" April/May, pp. 21, 27–32.
1985 "Opinion roundup: III—the exit poll results." December/January, pp. 28–29.

Rapoport, Ronald B.
1981 "The sex gap in political persuading: where the 'structuring principle' works." American Journal of Political Science 25:32–48.
1982 "Sex differences in attitude expression: a generational explanation." Public Opinion Quarterly 46:86–96.

Schneider, William
1984 "Opinion outlook: the democrats are counting on the gender gap, but it may not be much help." National Journal, June 23, 1984, pp. 1242–43.

Shabad, Goldie, and Kristi Anderson
1979 "Candidate evaluations by men and women." Public Opinion Quarterly 43:18–35.

Shapiro, Robert Y.
1982 "The dynamics of public opinion and public policy." Doctoral dissertation, University of Chicago.

Shapiro, Robert Y., and Benjamin I. Page
1984 "Subgroup trends in policy choices: a preliminary report on some theories and findings." Paper presented at the annual meeting of the Midwest Political Science Association.

Smeal, Eleanor
1984 Why and How Women Will Elect the Next President. New York: Harper and Row.

Smith, Tom W.
1978 "In search of house effects: a comparison of responses to various questions by different survey organizations." Public Opinion Quarterly 42:443–63.
1981 "Educated don't knows: an analysis of the relationship between education and item nonresponse." GSS Technical Report No. 28. NORC, University of Chicago.
1984 "The polls: gender and attitudes toward violence." Public Opinion Quarterly 48:384–96.

Stoper, Emily, and Roberta Ann Johnson
1977 "The weaker sex and the better half: the idea of women's moral superiority in the American feminist movement." Polity 10:192–217.

Stouffer, Samuel A.
1967 Communism, Conformity, and Civil Liberties. New York: Wiley. Original 1955.

Taylor, D. Garth.
1980 "Procedures for evaluating trends in public opinion." Public Opinion Quarterly 44:86–100.

Verba, Sidney, and Norman H. Nie
1972 Participation in America. New York: Harper and Row.

Welch, Susan
1977 "Women as political animals." American Journal of Political Science 21:711–30.

Wolfinger, Raymond E., and Steven J. Rosenstone
1980 Who Votes? New Haven, CT: Yale University Press.

Preelection Polls in Israel: Structural Constraints on Accuracy

JACOB SHAMIR

REPORTING preelection polls in the mass media is a popular and widespread practice in Western democracies. Much has been written on the definition of polls as news (e.g., Crespi, 1980; Tuchman, 1978), styles of reporting (Broh, 1980) and standards for accurate presentation of results (Field, 1971, Meyer, 1973; Noelle-Neuman, 1980; Miller and Hurd, 1982). Much of social science research methodology literature, including many of the articles that have appeared in this journal over the years, have devised techniques to improve the accuracy of such polls.

This article approaches the issue of election polls' accuracy from a different angle. It suggests that accuracy of reports on preelection polls is greatly affected by the environement in which pollsters operate as well as by structural characteristics of the system itself. Hennessy and Hennessy (1961) have already found that partisan-sponsored polls overestimate the actual vote for the candidate of the party for which the polls were done. They attribute this effect to pollsters' wish to please their partisan employer. However, their study suffered from serious sampling as well as other methodological problems (Perry,

---

Abstract This article examines the last two Israeli election campaigns of 1981 and 1984, and suggests that structural variables greatly affect the accuracy of reported preelection polls. Pollsters' quest for more detail and refinement driven by competition and mass media pressures for highly processed predictions is liable to result in less accurate reports. The kind of errors in polls' projections are situation-bound and greatly depend on the nature of the floating vote in the system. Interpollster accuracy, however, is mainly determined by pollster–sponsor dependencies. Pollsters working for political parties projected results more favorable to their sponsors or less favorable to their sponsors' opponents. Moreover, the stronger the pollster–sponsor structural dependencies, the greater the inaccuracy.

Jacob Shamir is Lecturer of Communication at the Journalism Program, Tel-Aviv University, Tel-Aviv, Israel.

Public Opinion Quarterly Vol. 50:62–75 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-62/$2.50

1962). I suggest that biasing pressures on pollsters are structurally inherent in the pollster–sponsor work relationship. The closer these relations get and the more structurally dependent pollsters become, the more favorable their reports are to their sponsors.

Structural characteristics of the system itself further affect accuracy and may facilitate such partisan pressures. Some political systems and situations are less transparent than others to preelection predictions. By transparency I mean the amount of practical and methodological difficulties that inherent systemic characteristics impose on pollsters' prediction efforts. A system's transparency depends on such factors as the number and size of parties in the system, the electoral system, and the public's willingness to volunteer valid voting intentions. The quest for higher precision and detail in a low transparency system—driven by competition and mass media pressures for highly processed predictions—is liable to result in less accurate reports, especially when partisan relationships with sponsors are involved. These propositions are examined in the Israeli system in the last two election campaigns of 1981 and 1984; however, they may be generalized to other polling situations and other political systems.

A System's Transparency to Polling Efforts

Israel's political system since the 1977 election, which marked the end of the dominant party system, poses a challenging environment to pollsters attempting to predict election outcomes with accuracy. Deep ideological cleavages coupled with a proportional representation electoral system resulted in two major political blocks almost equal in size. Each is composed of a major party and a host of satellite parties. A smaller third block consists of religious parties.

The abundance of political parties in the Israeli system, many of which are rather small, poses several methodological problems for pollsters. First, predictions with respect to the smaller parties, in particular those on the verge of passing the one percent barrier clause, are poor and risky because of the pollsters' commonly used sample size of 1200 subjects and its resulting sampling error. Second, it is much harder to classify no answer responses and undecided voters among the various political parties with high confidence. Subjects' stand on debated issues and attitudes toward party leaders may serve as powerful predictors in classifying floating voters among the two major blocks but are less valid in classifications among the various parties within a specific block or among other smaller parties. This finer voting distribution is critical in the Israeli context since it may determine the ability of the major parties to form a coalition government.

The major problem for Israeli pollsters, however, is the sizeable number of undecided voters, refusals to reveal voting intentions, and insincere responses in the system (Diskin, 1981). For example, voters of the first two kinds constituted about 30 percent of the population only one week before the 1981 and 1984 elections.[1] When the fractionalized party system interacts with the large number of ambiguous voting responses, the dimensions of the pollsters' problem increase many fold. A complex system of allocating residual Knesset seats among the various parties after the election poses another problem for pollsters attempting to cast their predictions in terms of number of seats.

These last two problems can be avoided if predictions are given in raw percentages rather than Knesset seats and hard to classify votes are left unclassified. The price, of course, is cruder prediction and less meaningful results for decision makers and mass media audience.

Additional problems are posed by two especially hard to reach populations of the Israeli constituency—soldiers and Arabs. These groups are usually not polled because of sampling, interviewing, and cooperation problems.

One may have expected a system such as this to encourage a conservative approach on the part of pollsters in treating the crude data obtained in the field and in the degree of specificity of their predictions. It seems, however, that with the increase in the public's interest in polls there is rather an increase in pollsters' attempts to provide highly processed data and more specific predictions. The number of pollsters casting their predictions in terms of seats won rather than percentages increased monotonically since 1977.

This trend correlates with changes in the amount and style of polls' coverage in the Israeli press. Weimann (1982), for example, found that polls' media coverage peaked at the 1981 election, reaching a state of "obsession for polls." The 1984 campaign was if anything even "worse" in this respect. Pollsters themselves became public figures heavily interviewed in the mass media and challenging each other's predictions.[2] The style of polls' coverage in the media reached the extreme characteristics of what Cantril (1976) and Broh (1980) called "horserace journalism." Thus, most reports on polls in 1984 focused on the size and dynamics of the gap between the two major parties, on the electoral fate of certain smaller extremist new parties, and on "beauty contest" ratings of the candidates for various government posts.

---

[1] Kessafim No. 54, 6/29/1981; Ma'ariv 7/19/1984.

[2] For example, Ma'ariv 7/13/1984, Hadashot 7/18/1984.

Data Refinement Efforts

Pollsters' main data refinement efforts center on developing valid classification methods to handle the great number of hard to classify responses obtained in the field. These methods, designed to refine predictions, may paradoxically become a major source of prediction error.

Commonly, undecided voters are investigated with respect to their acceptable alternative parties. The known alternatives are then used to classify them as either floating between the two major political blocks or within a major block. Then decision rules for final classifications take into account direct and projective attitudinal measures, subjects' declared vote in the last election, and at least some experience and judgment on the part of the pollster. Subjects who refuse to reveal their vote are handled much the same way except that it is more risky to initially classify them into one of the major blocks.

These procedures constitute a major source for classification errors for several reasons. The theoretical foundations of classification rules are quite weak and the boundaries between objective classification criteria and subjective judgment are not always clear, which opens the way to biasing structural effects. Second, these decision rules may also be very situation-bound, performing well in some respects, in one election, and proving to be wrong in another. Third, some of the objective classification measures may be of low validity. For example, subjects' information concerning their last vote may be false (Diskin and Felsenthal, 1981). Some pollsters handle this by using some aggregate data correction procedures which take into account the gap between subjects' declared voting distribution in the last election and the actual voting results.

Other potential sources of error are the assumptions made by pollsters on the vote distribution in the Arab sector. These assumptions are usually based on outdated last election statistics, subjectively adjusted to fit the present situation. As the Arab sector vote accounts for about 10 out of 120 Knesset seats, it may constitute a considerable source of prediction error. Since these data refinement procedures are based at least partly on subjective judgment and experience, it is easy to see how organizational, social, and economic factors may greatly affect their application.

Pollster–Sponsor Structural Dependencies

The last two elections, which varied considerably with respect to the pollster–sponsor relationships, can provide a good test for our propositions. In the 1981 campaign the Likud contracted one of the two largest commercial polling organizations in the country—Modi'in Ezrahi Research Institute (MERI). The Alignment in turn contracted the other one, which was affiliated with the advertising agency working with them—Dahaf. Both political parties, however, commissioned additional polls from other sources as well. Similarly, both Dahaf and MERI did not work exclusively for these political parties but for other parties as well, and provided regular polling services to several media organizations.

In 1984 the two major parties started by contracting commercial polling organizations which were dismissed later in the process in favor of free-lance pollsters operating ad hoc "in-house" polling facilities.[3]

Clearly, then, the Israeli polling industry witnessed considerable change in pollster–sponsor dependencies from the 1981 to the 1984 elections. First, free lance pollsters working exclusively for a political party have a considerably greater economic dependence than do polling organizations serving additional clients as well.[4] However, more important is the physical and social proximity stemming from the use of party facilities and the involvement of interested party officials in polling activities. Close proximity and ongoing meetings between the pollster and the sponsor's representatives are likely to reinforce informal social ties and the transmission of expectations concerning polls' results. Moreover, such conditions further decrease the pollster's control over the content and style of the poll's results, in a system already suffering from lack of self-regulation and public supervision over pollsters' activities.

The next section shows that systemic characteristics, i.e. the number and kind of the floating votes in the 1981 and 1984 Israeli elections, determined to a large extent the kind of errors in predicting election results. It also shows that pollsters working for political sponsors tend to favor their sponsors in preelection predictions. Moreover, inaccuracies in predictions increase with increases in pollsters' dependence on their political sponsors.

Analysis

The data used in the analysis consist of 31 out of 32 polls[5] carried out by Israel's most prominent pollsters and published in the press during

---

[3] The Alignment pollster established and directed a full-scale in-house polling organization, operating within party facilities. The Likud pollster also used party facilities to direct polling efforts, subcontracting outside research organizations for data collection.

[4] The Alignment pollster in 1984 seems to have volunteered, receiving only expenses from the party.

[5] See note c, Table 3.

Table 1. Styles of Reporting Preelection Projections in 1981 Israeli Election Campaign

	IASR[a] *(1st week of June)*			*SRC*[b] *(June 4–11)*	*Dahaf*[c] *(June 1–6)*	*MERI*[d] *(June 2–7)*
Party	*Type of Vote*	*% Projected*	*Party*	*% Projected*	*Seats Projected*	*Seats Projected*
Likud	definite	19	Likud	38	42	46
Alignment	definite	24	Alignment	33	43	40
Alignment or CRM/Shinuy/Shelli/Tami	floating	8[e]	NRP	6	6	8
Alignment or Telem/Tchiya/NRP	floating	3	Aguda		5	5
Alignment or Likud	floating	3	Telem		4	3
Likud or Telem/Tchiya/NRP/Tami	floating	10	Tami		4	3
Other parties	definite/floating	18	Tchiya		3	2
Unclassified	unknown	15[f]	CRM			2
			Other parties	9	13	11
			Undecided	14	15.8	14.5
			Not sure		19.2	

[a] Israel Institute for Applied Social Research.
[b] Smith Research Center.
[c] Dahaf Research Institute.
[d] Modi'in Ezrahi Research Institute.
[e] Indicated vote preference but not sure in their decision.
[f] Did not indicate vote preference (don't know or no answer).

the last two months of the 1981 and 1984 election campaigns. These are the hottest months in terms of political parties' struggle for voters' support and the most crowded with preelection predictions appearing in the mass media.

Most polls were published on a regular basis under various contractual arrangements between the Israeli major newspapers and their pollsters. Some polls, however, in particular the Likud and Alignment polls in 1984, were made public either in press conferences or less formally through political correspondents covering the election campaign.

Let me first illustrate the various pollsters' approaches in reporting their results. Table 1 presents results of four polls carried out by four prominent Israeli polling organizations less than a month before 1981 election day. The results are presented in the same format published in the press.

The first major point evident from the table is the sizeable percentage of respondents floating between at least two alternative parties. IASR, which is a nonprofit research organization affiliated with the Hebrew University in Jerusalem, reports floaters' data in great detail. Only 43 percent of the population is unequivocal with respect to their preferred party—the Likud or the Alignment. Most others either float between the two major parties (Likud and Alignment), between one of those and smaller satellite parties, or do not provide any indication whatsoever on their intended vote.

Table 1 also shows that Israeli pollsters approach this problem with varying degrees of conservatism. The most conservative approach is IASR's approach, which strictly avoids any conclusive classification of floating voters. Instead, it reports the raw percentages obtained in the field, for all possible combinations of voting alternatives considered by subjects. SRC, whose director regularly provides the Israeli TV election night projections, takes a less conservative approach. Although results are still reported in percentages, it seems that all floating votes, except for 15 percent hard to classify subjects, were allocated to one of the party blocks. This classification attempt was greatly simplified by collapsing the fractionalized party system into four major blocks: Likud, Alignment, religious parties, and others.

A much more daring approach is taken by Dahaf and MERI, which are the largest commercial polling organizations in the country. In addition to classifying floaters, both report projections in Knesset seat numbers, present a more detailed list of political parties, and a smaller "other" waste basket category. This indicates elaborate classification schemes on their part allowing relatively finer classification decisions.

However, with increased pressures on the part of the mass media to provide more highly processed and more exciting data to the public, we notice an interesting shift in some of these reporting styles. Subsequent IASR reports in 1981 adopt SRC reporting style and present results

Table 2. Final Projections of Election Results by Israeli Pollsters in 1981 and 1984, and Mean Deviations from Actual Outcomes

	1981				*1984*					
Party	*MERI 6/24–25*	*SRC 6/24–28*	*Dahaf 6/29*	*Election Results 6/30*	*LP*[a] *7/11*	*AP*[b] *7/15*	*Dahaf 7/18*	*MERI 7/19*	*SRC 7/19*	*Election Results 7/23*
Likud	42	43(36%)[c]	43	48	40	36	37	40	41	41
Alignment	42	42(35%)	43	47	45	52	49	46	48	44
L + A	84	85	86	95	85	88	86	86	89	85
L − A	0	1	0	1	−5	−16	−12	−6	−7	−3
Other	36	35(29%)	34	25	35	32	34	34	31	35
A and L overall deviation			5.03					3.91		
L + A deviation			10.03					2.32		
L − A deviation			.82					7.47		

[a] Likud pollster.

[b] Alignment pollster.

[c] The percentages originally reported by the pollsters were translated by this author into seat numbers ($N = 120$) for better comparability and for the purpose of calculating standard deviations.

after a more conclusive classification of floaters. SRC in turn shifted from exclusive percentage-based predictions toward predicting number of seats at the end of the 1984 campaign.

I turn now from the styles of reporting polls which reflect pollsters' refinement efforts to the major issue in this article—accuracy. I define accuracy as the degree of congruence between pollsters' final preelection predictions and actual results. However, for polls conducted earlier in the campaign interpollster consistency may serve as another accuracy measure.

Table 2 relates to the first meaning of accuracy. It presents final predictions of three pollsters in 1981 and five pollsters in 1984, as well as actual election outcomes. The results were collapsed into three categories (Likud, Alignment, and All Others). Two additional rows present the major parties' total number of seats combined (L + A) and the seat difference between them (L − A). The last rows present several accuracy measures. The overall error row presents mean deviations of pollsters' predictions of each of the major parties—from actual results.[6] The L + A error row presents deviations across pollsters of the combined number of seats predicted for the major parties from actual results.[7] Similarly the L − A error row presents mean error predictions of seat difference between the two major parties.[8]

---

[6] The computational formula was:

$$\mathrm{ME} = \sqrt{\sum \frac{(X_{ij} - A_i)^2}{n}}$$

where ME = mean error
i = 1,2 parties
j = pollsters
X_{ij} = last election prediction of pollster j for party i
A_i = actual results for party i
n = total number of predictions.

The "all other" category was excluded in the calculation for convenience reasons since it complements the major parties' number of seats to 120—the number of seats in the Israeli Knesset.

[7] The computational formula is similar to the one presented in footnote 6:

$$\mathrm{MES} = \sqrt{\sum \frac{(X_j - a)^2}{n}}$$

where MES = mean error sum
j = pollsters
n = total number of predictions
X_j = last election prediction of pollster j for (L + A)
a = actual ($L + A$) result.

[8] The computational formula is the same as in footnote 7, replacing MES with MED (mean error difference) where X_j is the last election prediction of pollster j for ($L - A$) and a is the actual ($L - A$) result.

Examining pollsters' overall error in 1981 and 1984, we see that although the 1981 error is larger, the two are not very different. In 1981 MERI, SRC and Dahaf deviated by an average of about five seats in their predictions for the two parties. In 1984 the Likud pollster (LP), the Alignment pollster (AP), Dahaf, MERI, and SRC deviated by an average of about four seats. However, it is evident that the source of these errors is different in the two elections. The major source of error in 1981 lies in inaccurate predictions of the combined number of seats of the two major parties (mean deviation of 10.03), while the difference between the two major parties was accurately predicted (.82 mean deviation). In 1984 the picture is reversed. Errors in the predicted sum number of seats are small (2.32), while most errors lie in the predicted difference between the major parties (7.47). What accounts for these differences? The answer lies in different misclassification patterns stemming from differences in the nature of the floating vote in 1981 and 1984.

From the IASR data in Table 1, it is clear that pollsters' major dilemma in 1981 was to classify the sizeable percentage of voters floating among each major party and its satellites. Many fewer voters floated between the two major parties themselves (Diskin, 1981), accounting for a negligible 3 percent in the IASR data. In light of the close run between the major parties, it seems that there was no justification to apply different classification rules for the Likud and the Alignment floaters. The sizeable error in predicting the total number of seats for the major parties indicates that pollsters consistently erred in allocating more floaters to the smaller rather than the major parties.[9]

The 1984 election differed from 1981 in two respects. First, the competition was not as close as in the previous election. Secondly, many more voters floated between the two major parties. A national poll carried out by the end of June 1984[10] showed that floaters between each major party and its satellites account for 9 percent of the total vote compared to about 18 percent in 1981. However, more than 10 percent of the total vote floated between the Likud and the Alignment, compared to 3 percent in 1981. Although this last percentage seems to have dropped as election came closer, the overall number of floaters between the two major parties was considerably greater than in 1981. In 1984, then, more classification errors occurred between the two major parties and fewer within each party block than in the previous election

---

[9] Some pollsters argued that the polls' predictions themselves showing a close run motivated floaters to lean toward the major parties in their final decision. However, since a strong theory of floaters' classification should be able to account for this possibility as well, we may still consider it a misclassification error.

[10] Dahaf poll.

Table 3. Israeli Polls Results and Interpollster Consistency Measures in the Last Month Before the Election: 1981 and 1984

Pollster Data Collection: 1981

	Dahaf 6/1–2	*MERI* 6/2–7	*SRC*[a] 6/4–11	*IASR*[b] Beg. June	*MERI* 6/15–18	*Dahaf* 6/16–17	*IASR* 6/14–18	*Dahaf*[c] 6/26	*MERI* 6/24–25	*SRC* 6/24–28	*Dahaf* 6/29	*Elect. Results* 6/30
Likud	42	46	45(38%)	19%	49	45	47(39%)	40	42	43(36%)	43	48
Alignment	43	40	40(33%)	24%	37	38	38(32%)	43	42	42(35%)	43	47
L + A	85	86	85		86	83	85	83	84	85	86	95
L − A	−1	6	5		12	7	9	−3	0	1	0	1
Other	35	34	35(29%)	57%	34	37	35	37	36	35(29%)	34	25
S.d. (L + A)			.48			1.25			.5	.5		
S.d. (L − A)			3.09			2.06			1.5	.5		

Pollster Data Collection: 1984

	Dahaf 7/1–3	*MERI* 7/1–4	*LP* 7/1–4	*AP* 7/8	*SRC* 7/8–10	*Dahaf* 7/11	*MERI* 7/11–12	*LP* 7/11	*AP* 7/15	*Dahaf* 7/18	*MERI* 7/19	*SRC* 7/19	*Elect. Results* 7/23
Likud	38	37	41	37	35(29.5%)	41	34	40	36	37	40	41	41
Alignment	53	47	48	53	47(39.5%)	50	48	45	52	49	46	48	44
L + A	91	84	89	90	82	91	82	85	88	86	86	89	85
L − A	−15	−10	−7	−16	−12	−9	−14	−5	−16	−12	−6	−7	−3
Other	29	36	31	30	38	29	38	35	32	34	34	31	35
S.d. (L + A)		2.69					3.49				1.41		
S.d. (L − A)		3.67					3.92				2.62		

[a] The percentages originally reported by the pollsters were translated by the author into seat numbers ($N = 120$) for better comparability and for the purpose of calculating standard deviations.

[b] The IASR poll is not comparable to other polls in its group as floating voters remained unclassified. Thus, it was not taken into account in calculating standard deviations.

[c] Another Dahaf poll dated 6/23–24 was excluded from this group since it did not account for the TV debate between the major party leaders as well as a personality change in the alignment leadership—occurring shortly afterwards. The remaining polls in the group accounted in part or fully for these events. The results of this Dahaf poll are very close to the previous Dahaf poll from the 6/16–17.

year. This, of course, explains at least partly the greater accuracy in the sum number of seats for the major parties compared to their projected seat difference.

Table 3 presents 23 polls of major Israeli pollsters published in the mass media in the last month before the elections in 1981 and 1984. The published results are grouped according to polls' date of execution,[11] and collapsed as in Table 2.

If we are willing to accept final results as an accuracy standard for earlier polls as well[12] we can see the same error pattern occurring all along. When we examine the (L + A) and (L − A) rows in Table 3 it is clear that all 1981 polls consistently underestimated the major parties' combined number of seats, but quite accurately predicted the difference between them. In 1984, however, all polls overestimated the difference between the two major parties (L − A), whereas the combined number of seats (L + A) predictions were close to the final results all along.

---

[11] Eight additional polls carried out during May 1981 and June 1984—two months before election date—are not presented here for space reasons. However, they are taken into account in subsequent analysis. In classifying the various polls into groups, we made sure that major political events with a potential effect on public opinion do not separate polls classified in the same group.

[12] This is a problematic assumption suggesting no overtime effect on pollsters' overall prediction accuracy. Testing for this possibility we found that a one-way analysis of variance failed to show significant accuracy differences when contrasting latest predictions with all other predictions. In other words, later clusters did not show significantly greater overall accuracy compared to earlier ones, neither in 1981 nor in 1984.

Turning to the second definition of accuracy, we examined interpollster predictions' consistency within each of the 1981 and 1984 polls clusters. This was done by simply calculating for each cluster the standard deviation from the mean number of seats predicted for the two major parties combined (L + A) and from the mean seat difference predictions (L − A). These deviations are presented in rows six and seven of Table 3.

In contrast to the previous accuracy measure, we do not observe here any interaction of the sum and difference deviations by election year. Instead, standard deviations of (L − A) in the last month before the election are consistently greater than (L + A) deviations, in both years.[13] I suggest that the difference in results evident in the two accuracy measures stems from their sensitivity to different systemic error sources. While deviations from actual results as an accuracy measure reflect floating vote allocation problems, interpollster consistency mainly taps pollster–sponsor dependencies as a source of variability. These dependencies are best reflected in the (L − A) measure, since both politicians and mass media consider the seat difference to be a highly effective cue on campaign momentum in the eyes of the public. Biasing pressures will thus affect the seat difference predictions more than other possible measures, causing greater variability among pollsters on it.

This can be easily seen when predicted seat differences between the major parties are examined in greater detail. For example, Dahaf, working for the Alignment, predicted at the beginning of June 1981 a one-seat difference in favor of its client, while MERI, serving the Likud, predicted its client to lead by a margin of six seats. In the next polling wave, MERI projected an increasing 12-seat gap in favor of the Likud, while Dahaf admitted only a 7-seat difference. By the end of the month competition became close again and was reflected in the 42-seat tie projected by MERI. Dahaf in turn predicted a 3-seat lead for the Alignment. Dahaf's final prediction on June 29 converged to a 43-seat tie while the independent SRC predicted the Likud to lead by one seat. Clearly, then, the systematically biased Dahaf and MERI predictions in favor of their respective sponsors constitute the major source of interpollster variability within time clusters.

The 1984 pattern is no different in this respect. This time we compare the Alignment and Likud free-lance pollsters (AP and LP, respectively). At the beginning of July the Alignment pollster projected a 16-

---

[13] Earlier polls, however—the May polls in 1981 and the June polls in 1984—are not consistent with this observation.

Table 4. Degree of Dependency by Absolute Deviations of (L − A) Predictions from Clusters' Mean for 29 Israeli Polls

Interpollster Accuracy	*Dependency*		
	Low	*Medium*	*High*
0.2 –2.2	71	50	20
2.21–4.2	29	40	20
4.21–6.2	0	10	60
	100%	100%	100%
	(14)	(10)	(5)

$\chi^2 = 12.1$(4 d.f.), $p < .05$ Gamma = .60

seat difference in favor of its client, compared to only a 7-seat difference projected by the Likud's pollster. By mid July the Alignment pollster still stuck to this 16-seat difference while the Likud pollster moved down even further to a narrow 5-seat difference. These mid July projections were also the final AP and LP official predictions released to the press. Other pollsters' final predictions of seat differences all fell between these extremes.

The pattern is very clear—pollsters who were structurally dependent on one of the major political parties favored their sponsor both in 1981 and 1984 election campaigns. With no exception, this pattern also characterizes eight additional polls carried out one month earlier in both campaign years (these polls are not presented in the table for space reasons). In all cases involving partisan relationships, pollsters tended to project results more favorable to their clients or less favorable to their clients' opponents. However, for a more complete support of our hypothesis it is possible to restate the findings of Table 3 somewhat differently, to show the existence of an ordinal relationship between the degree of dependency and accuracy. This is presented in Table 4.

In 1984, the Likud pollster (LP) and the Alignment pollster (AP) worked under highly dependent conditions, operating from sponsors' facilities and working exclusively for their political sponsor. Medium dependency conditions characterized Dahaf's and MERI's environment in 1981. Both operated from their own long established facilities and were commissioned by media organizations in addition to their major political sponsors. SRC's June poll in 1984 (not presented in Table 3) fell into the same category.[14] Totally independent in these respects were SRC in 1981 and Dahaf and MERI in 1984. They did not work for any of the major political parties in the last two months before the election. SRC's July polls in 1984 fit into this category as well.

---

[14] SRC's June poll was carried out still under contract with the Alignment. Soon afterwards SRC stopped working for the Alignment, which retained and enlarged the in-house polling enterprise headed by a university professor on leave.

Table 4 shows the percentage of high, medium, and low accuracy polls in each of the three dependency categories. It is based on 29 polls including those presented in Table 3 and those conducted one month earlier in both election years.[15] Accuracy of each specific poll was measured as its (L − A) absolute deviation from its group's (L − A) mean. The smaller the distance between the cluster's mean and the specific poll's prediction, the greater the accuracy. It is clear that the greater the pollster's dependency, the less accurate are his seat difference predictions. Only one out of the five polls conducted by highly dependent pollsters reported parties' seat difference accurately. The results of Table 3 do not leave any doubt as to the direction of bias. Fifty percent of the medium-dependent pollsters' polls were accurate in this sense, while more than 71 percent of independent pollsters' polls were highly accurate.

Discussion

Our findings suggest that processes affecting accuracy in partisan-sponsored polls are more complex than the Hennessy and Hennessy (1961) contention that pollsters wish to please their partisan employers. The fact that pollsters' tendency to favor their sponsors varies monotonically with the strength of their structural dependencies points to the role of structural factors affecting accuracy. For example, it is likely that prolonged hours of physical proximity to sponsors' campaign facilities and close social ties with campaign officials expose pollsters to high levels of optimistic expectations as to the campaign outcome. Thus, the closer the relationships, the stronger the biasing social pressures directed toward pollsters. Under such conditions it is perhaps easier to interpret vague data in a favorable rather than an unfavorable direction to surrounding expectations. Although not directly tested in this study, we suspect that pollsters' data refinement efforts in a low transparency system, based partly on subjective judgment and experience, are not rigid enough to withstand environmentally biasing pressures.

The role of pollsters' data refinement efforts as a source of error becomes more evident when final election predictions are examined. The different nature of pollsters' prediction errors across election years can be partly attributed to classification errors of the sizeable floating vote in the system. Thus, attempting to provide greater detail and refinement based on relatively weak theories in a low transparency system, researchers expose themselves to greater risk of error. This raises an important normative question with respect to pollsters' role in election times. Clearly, in the Israeli context pollsters were driven both by competition and by a "horserace" attitude on the part of the press to provide highly processed projections that are also more exciting. Their failure resulted in a decline in status and prestige of the entire polling industry. An alternative course that seems more promising, although less glossy, is to focus on descriptive and explanatory aspects of the data as long as existing behavioral theories do not provide a more sound base for accurate predictions.

---

[15] Two IASR polls were not included in the analysis as they did not attempt to classify the floating votes and are thus not comparable to the rest.

References

Arian, A.
1981 "Elections 1981: competitiveness and polarization." Jerusalem Quarterly 21:3–27.
Broh, A. C.
1980 "Horse race journalism: reporting the polls in the 1976 presidential election." Public Opinion Quarterly 44:514–29.
Cantril, A. H.
1976 "The press and the pollster." Annals of the Academy of Political and Social Science 427:45–52.
Crespi, I.
1980 "Polls as journalism." Public Opinion Quarterly 44:462–76.
Diskin, A.
1981 "The 1981 election: public opinion polls." Jerusalem Quarterly 22:99–104.
Diskin, A., and D. S. Felsenthal
1981 "Do they lie?" International Political Science Review 2:407–22.
Field, M. D.
1971 "Towards responsibility in reporting opinion surveys, the researchers' view." Public Opinion Quarterly 35:342–46.
Hennessy, B. C., and E. R. Hennessy
1961 "The prediction of close elections: comments on some 1960 polls." Public Opinion Quarterly 25:405–11.
Meyer P.
1973 Precision Journalism, A Reporter's Introduction to Social Science Methods. Bloomington: Indiana University Press.
Miller, M. M., and R. Hurd
1982 "Conformity to AAPOR standards in newspaper reporting of public opinion polls." Public Opinion Quarterly 46:243–49.
Noelle-Neuman, E.
1980 "The public opinion research correspondent." Public Opinion Quarterly 44:585–97.
Perry, P. K.
1962 "Correspondence." Public Opinion Quarterly 26:133–35.
Tuchman, G.
1978 Making News: A study in the Construction of Reality. New York: The Free Press.
Weimann, G.
1982 "Press Coverage of pre-election polls in Israel." State, Government and International Relations 19–20:132–43 (Hebrew).

Mass Communication and Political Socialization: Specifying the Effects

GINA M. GARRAMONE AND CHARLES K. ATKIN

STUDIES examining the role of mass communication in young people's learning about politics indicate that the media contribute to political socialization, but the research does little to unravel the complexities of the process. By treating the media exposure patterns, the criterion variables, and the audience as unidimensional concepts, these studies fail to tease out the more specific effects that might exist for *certain* media, *certain* variables, or *certain* audience categories. The present study attempts to address these shortcomings by comparing the socialization effects of exposure to four media sources of political information—television news, radio news, newspaper news, and newsmagazines—on several specific types of political knowledge and behavior. The study also explores differences in effects for certain subgroups of young people.

Abstract A survey investigation of teenagers examined differential socialization effects of four news media on four types of political knowledge and behavior among several categories of youth on the basis of abilities and predispositions. Regression analyses of questionnaire responses from 280 seventh and tenth grade students show that TV news exposure is the strongest predictor variable. The broadcast media have a greater impact on current events knowledge than fundamental political knowledge, while a slight reversal occurs for the print media. Newspaper reading has the greatest effect on participation. There are predominantly uniform relationships across subgroups, providing little evidence that the news media produce knowledge or behavior gaps.

Gina Garramone is Assistant Professor of Advertising and Charles Atkin is Professor of Communication and Telecommunication, both at Michigan State University. The research was supported by the National Association of Broadcasters.

Public Opinion Quarterly Vol. 50:76–86 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-76/$2.50

Four Types of Media

Although entertainment and persuasive content in the mass media contain some politically relevant messages, news items provide the bulk of political information. A number of studies demonstrate widespread consumption of hard news by young people (Gollin and Anderson, 1980; Drew and Reeves, 1980a; Atkin, 1978). The inputs from news media usage do not constitute a unidimensional phenomenon, however. The four news media differ in terms of inherent *channel mode* attributes, conventional *style* of news presentation, and nature of *content*.

The television medium can present visual images, while the other media are largely restricted to verbal modes. The print media allow considerable receiver control of information selection and processing, requiring some degree of motivated and focused attention. The sender-controlled broadcast transmissions require minimum audience effort (Chaffee and Tims, 1982).

Stylistically, the print news media tend to feature greater complexity with respect to technical readability and substantive intricacy; broadcast news is presented in a simplified and concrete style that facilitates comprehension. In addition, the dramatic, personalized, and entertaining presentational style of television news is highly engaging relative to the other media.

While the topical content differs minimally across media, newspapers convey a larger quantity of messages, and print news items are characterized by greater depth coverage, with fuller detail, broader perspective, and richer background information compared to the ahistorical and abbreviated broadcast content.

Each of these dimensions has implications for differential impact on adolescents at the cognitive and behavioral levels.

Political Criterion Variables

KNOWLEDGE

Past studies have found a relationship between media exposure and young persons' current events knowledge (Chaffee et al., 1970; Atkin and Gantz, 1978) and more fundamental knowledge about the political scene (Conway et al., 1975; Gollin and Anderson, 1980). *Current events knowledge* comprises knowledge of the discrete, time-bound events and associated individuals featured in the news media. In the 1981 period of this study, for example, IRA hunger-striker Bobby Sands died. Such concrete, specific events and the names of the individuals associated with them typically receive heavy media coverage. These events are not independent of the political history and extant political milieu in the respective country, of course. The historical conflict in Northern Ireland, including the religious and political orientations of the combatants, provides a certain perspective on the hunger strikers. Such *fundamental knowledge*, of political philosophies, structures, and historically significant developments, constitutes a more basic understanding underlying daily political events.

Most studies have limited their analyses to only one type of knowledge and only one type of medium, precluding comparisons of differential effects. On the basis of the mode, style, and content differences between broadcast and print media, however, we may predict differential effects likely to emerge. Television as a news medium, for example, compels contact to each story, facilitating exposure to a variety of topics. Compared to print, broadcast news stories tend to be briefer, simpler, easier to process, and more engaging in presentational style. All of these factors suggest that the broadcast media, and television in particular, should be especially effectual for young audiences, who tend to be visually oriented and less cognitively sophisticated than adults. We predict, therefore, that broadcast news exposure will be more strongly related to overall political knowledge than will print news exposure.

Chaffee and Tims (1982) noted that exposure to the superficial broadcast media may be particularly effective for learning such "itemized" cognitions as the names of political actors or parties. Regular use of print media, on the other hand, with their depth of coverage, substantive complexity, and self-paced processing, may be more likely to enable the youngster to acquire and maintain complex cognitive associations between the various elements of the political scene. We predict, therefore, that exposure to broadcast news will contribute more to knowledge regarding current events than to fundamental knowledge, ~~while~~ print news exposure will have greater impact upon fundamental than on current events knowledge.

BEHAVIOR

Although youngsters cannot participate formally in the political system, they can practice interpersonal discussion behaviors and form future participatory dispositions. The news media provide raw information that can be conversationally shared with others (Atkin and Gantz, 1978; Drew and Reeves, 1980a). By providing youngsters with information regarding others' political behavior, mass media exposure also may contribute to the development of long-range predispositions regarding such anticipated political activities as voting and campaign work (Sheinkopf, 1973). The mode, style, and content characteristics of the broadcast and print classes of media may also determine their differential impact on the two types of political behavior. The broadcast media, with their dramatic and entertaining presentation style, should have greater impact on interpersonal discussion than on the more complex anticipated participation. The more involving and educative print media, with their detail and broader perspective, should be more strongly related to anticipated participation than to discussion.

Categories of Young People

The various capabilities and orientations of the young media audience will influence media effects. For example, teenagers possessing greater cognitive sophistication are predicted to acquire complex political information more broadly and deeply than do younger children (Atkin and Gantz, 1978; Drew and Reeves, 1980b). The impact of cognitive sophistication on learning from the media may also depend upon the type of medium under consideration; younger and older students should learn equally well from the *broadcast* media, but older students should outperform younger students in learning from the more complex *print* media. The organismic variable of age is most often used as an indicator of cognitive sophistication. A child's scholarly orientation (e.g., grades in school, intention to attend college) is also likely related.

Interest in politics is another significant intervening variable. Youngsters with greater interest should more readily respond to news coverage. The impact of interest, however, may depend upon the class of media under consideration. Interest may have less impact on learning from the less involving broadcast media than from print media.

Finally, antecedent variables which may shape the political learning process are assessed. Research indicates that boys are more attuned to political news and more involved politically than are girls, and that socioeconomic status is slightly related to these variables (Atkin, 1981).

In summary, it is predicted that broadcast news exposure will be more strongly related to overall political knowledge than will print news exposure. It is also proposed that broadcast news exposure will be a better predictor of current events knowledge than of fundamental knowledge, but that print news exposure will be a better predictor of fundamental than of current events knowledge. Similarly, it is predicted that broadcast exposure will be more closely related to interpersonal discussion than to anticipated participation, but that print exposure will be more closely related to anticipated participation than to discussion. In addition, the relationships between media exposure and the criterion variables are expected to vary for subgroups of the young audience at various levels of age, scholarly orientation, and political interest.

Method

Questionnaires were administered to public school students in two mid-sized midwestern cities with typical socioeconomic characteristics. Students were drawn from the seventh grade ($N = 149$) and tenth grade ($N = 131$) classrooms at each site. The questionnaire contained items measuring the four types of media exposure, political knowledge and behavior, age, scholarly orientation, political interest, gender, and socioeconomic status. The variables were operationalized in the following manner:

Television News Exposure. The index of exposure to television news summed items dealing with how often national network newscasts were viewed, how many days had been watched the previous week, whether the news was viewed the day before the survey, the number of days viewed in an average week, and the proportion of the program that was usually watched; in addition, the frequency of viewing morning news-variety programs was included.[1]

Radio News Exposure. The radio news listening index summed the estimated number of radio news shows heard on an average day, on the previous day, and the amount of attention paid to radio news stories.

Newspaper News Exposure. This index summed respondents' answers to how often they read the front page of the newspaper and how many hard news articles they read per day.

Newsmagazine Exposure. This was measured by frequency of reading the weekly periodicals such as *Time* and *Newsweek*.

Some analyses call for a comparison of the two classes of broadcast and print media. For these analyses, television news exposure and radio news exposure are summed to form a *broadcast exposure* index. Newspaper news exposure and newsmagazine exposure are summed to form a *print exposure* index.

Political Knowledge. An index for *overall* political knowledge was constructed by summing together 46 questionnaire items, most of the "multiple choice" variety. The typical item posed a question about events, developments, or people in the political sphere, and listed three choices and the fourth alternative of "??" which students were encouraged to use if they were unsure.[2] This overall knowledge index is used in the contingent correlation analysis.

---

[1] The means, standard deviations, and Cronbach's alpha reliability coefficients for the indices are as follows: television news exposure, 17.19, 8.35, .78; radio news exposure, 6.31, 2.82, .71; newspaper news exposure, 4.72, 1.69, .62; newsmagazine exposure, 1.62, .67, estimated at .50; broadcast exposure, 23.50, 9.48, .78; print exposure, 6.34, 1.97, .63; overall political knowledge, 70.47, 8.64, .89; fundamental knowledge, 10.93, 2.08, .68; current events knowledge, 12.88, 2.34, .69; interpersonal discussion, 16.69, 5.04, .88; and participation, 6.32, 2.04, .51.

From the 46 knowledge items, the two more specific knowledge indices were constructed. The *fundamental* knowledge index was created by summing together eight items concerning historical facts or trends relevant to more current events.[3] To make a *current events* knowledge index comparable to the fundamental knowledge index, eight current news items were selected to match as closely as possible those measured in the fundamental knowledge index.[4]

Political Behavior. The *interpersonal discussion* index is composed of eight items dealing with interpersonal communication with both family and friends about four news topics.[5] The *participation* index measures anticipated election campaign activities, inclination to run for office, and plans to vote in elections.

Contingent Variables. Three variables were measured for contingent analyses: age, scholarly orientation, and political interest. *Age* is indicated by the student's grade in school. *Scholarly orientation* is measured by an index summing self-reported performance in school classes and intention to attend college. The *political interest* index measures how interested the student is in several news topics.[6] For use in the contingent analyses, each of these variables was divided at the median.

---

[2] Sample knowledge items include: "President Reagan recently made a trip to give a speech, which was the first time he left Washington since he was wounded. Where did he travel to?" "The leader of Japan visited Washington to talk with President Reagan. What did they talk about?" "What is the name of the man who is Budget Director for the U.S.?" "A few months ago, the 52 hostages came back to the United States. In what country were the hostages held prisoner?" "How many hunger strikers have died in Northern Ireland prisons?" "There have been debates about whether the U.S. should send military aid to a Central American country. Which country?" "Which country just elected a new president?" "What does Ronald Reagan want to sell to Saudi Arabia?"

[3] Sample items are: "Poland is run by a Communist government," "The conflict in Northern Ireland is caused by problems between two religious groups."

[4] Sample items are: "What are the people of Poland worried about this year?" "A man in Ireland named Bobby Sands died several weeks ago. How did he die?"

[5] The four topics are: President Reagan, things that happen in other countries (Poland, Russia, Israel, Iran, El Salvador are listed as examples), government and things that happen in Washington, and current problems in the news (crime, inflation, nuclear power, gun control are listed as examples).

[6] The topics included "things that happen in other countries," "things that President Reagan does," "government and things happening in Washington," and politics in general.

Table 1. Zero-order and Corrected Correlations Between Media Exposure and the Criterion Variables.[a]

	Knowledge			*Behavior*	
	Overall	*Current Events*	*Fundamental*	*Interpersonal Discussion*	*Anticipated Participation*
Broadcast	.43 (.52)	.39 (.53)	.26 (.35)[b]	.45 (.54)	.29 (.46)[c]
Print	.32 (.43)[d]	.24 (.36)	.24 (.36)	.39 (.52)	.37 (.65)[e]

[a] The correlations corrected for attenuation (Nunnally, 1967) are represented within parentheses. Corrections are made based on the internal consistency estimates (Cronbach's alpha) for each index.

[b] For Broadcast exposure, the difference between Current Events versus Fundamental knowledge attenuation-corrected correlations is significant at the 5 percent level.

[c] For Broadcast exposure, the difference between Interpersonal Discussion versus Anticipated Participation attenuation-corrected correlations is significant at the 5 percent level.

[d] For Overall Knowledge, the difference between Broadcast versus Print exposure attenuation-corrected correlations is significant at the 5 percent level.

[e] For Print exposure, the difference between Interpersonal Discussion versus Anticipated Participation attenuation-corrected correlations is significant at the 5 percent level.

Background Variables. In addition to the contingent variables, the student's gender and family socioeconomic status[7] were measured for control purposes. In some analyses, age, scholarly orientation, and political interest are also treated as background predictor variables.

ANALYSIS

Correlation and regression analysis are used to explore the relationship between media exposure and each of the criterion variables. Contingent correlations between media exposure and the criterion variables at high versus low levels of age (grade), scholarly orientation, and political interest also are computed.

Results

The results are described in the following order: (1) the relationship between media exposure and political knowledge, (2) the relationship between exposure and political behavior, (3) the contingent analyses.

POLITICAL KNOWLEDGE

Zero-order correlations were computed between broadcast and print classes of media exposure and overall political knowledge, current events knowledge, and fundamental knowledge. Because of the different reliabilities of the indices, the correlations are corrected for attenu-

---

[7] Socioeconomic status is indexed by parental occupation, social class of neighborhood, and family member college education.

Table 2. Regression of Criterion Variables on Demographic and Media Variables

	Knowledge			*Behavior*	
	Overall	*Current Events*	*Fundamental*	*Interpersonal Discussion*	*Anticipated Participation*
SES	.08[a]	.06	.08	−.04	.06
Gender[b]	.25	.26	.13	.06	−.06
Grade[c]	.27	.28	.27	.03	−.16
Scholarly orientation	.39	.28	.31	.38	.33
(R^2 for step)	(.33)	(.25)	(.23)	(.14)	(.14)
Television	.27[d]	.27	.16	.25	.08
Radio	.07	.10	.02	.14	.08
Newspaper	.04	.02	.05	.19	.22
Newsmagazine	.09	.04	.09	.04	.07
(Incremental R^2 for step)	(.11)	(.09)	(.04)	(.17)	(.09)

[a] Coefficients in the first four rows are standardized beta weights controlling for other demographic variables; $\beta > .11$ significant at 5 percent level.

[b] Gender is coded female-male.

[c] Grade in school is coded 7th-10th.

[d] Coefficients in the bottom rows are standardized beta weights controlling for demographic variables and other exposure variables; $\beta > .11$ significant at 5 percent level.

ation (Nunnally, 1967) using the internal consistency estimates. Both sets of correlations are presented in Table 1. The description of results and tests of significance pertain to the attenuation-corrected correlations. As expected, broadcast news exposure is more strongly correlated with overall political knowledge than is print news exposure (t (3,277) = 2.78, $p < .05$). While broadcast exposure also is strongly correlated with current events knowledge, it is significantly less strongly correlated with fundamental knowledge (t (3,277) = 4.03, $p < .05$). For print news exposure, the correlations with current events and fundamental knowledge are identical.

Hierarchical regressions were computed to provide a more detailed analysis of the relationship between the media exposure and knowledge variables, with background variables controlled. Entered in the first step of the regressions were socioeconomic status, gender, grade in school, and scholarly orientation (Table 2). The four media exposure indices were added as a block for the second step.

Three of the four background variables are significantly related to overall political knowledge. Males, tenth graders, and students with a strong scholarly orientation are more knowledgeable politically. Of the four media exposure variables, the only significant predictor is television news viewing, with a standardized beta coefficient (β) of .27. The next strongest predictor is newsmagazine reading.

Corresponding regression analyses were computed for current events and fundamental knowledge. As a block, the four media variables explain twice as much variance in current events knowledge ($IR^2 = .09$) as in fundamental knowledge ($IR^2 = .04$). Television news viewing is the best predictor of both current events and fundamental knowledge, but the relationship is much stronger for the former. Radio news listening is also a stronger predictor of currents events knowledge than of fundamental knowledge. Both of the print media indices, however, are slightly better predictors of fundamental knowledge than of current events knowledge. Addition of the political interest variable into the knowledge regression equations has little impact on the media coefficients.

POLITICAL BEHAVIOR

Zero-order and attenuation-corrected correlations were computed between media exposure and interpersonal discussion and anticipated participation (Table 1). Again, the description of results and tests of significance pertain to the attenuation-corrected correlations. The predicted pattern of differences between the discussion and participation correlations emerges: broadcast news exposure is more strongly related to interpersonal discussion than to anticipated participation ($t\ (3{,}277) = 1.66$, $p < .05$), and print news exposure is more strongly related to anticipated participation than to discussion ($t\ (3{,}277) = 3.07$, $p < .05$).

Hierarchical regressions were computed to assess the relationship between media exposure and the political behavior indices, with background variables entered first (Table 2). Students with a strong scholarly orientation are significantly more likely both to engage in interpersonal discussion regarding politics and to anticipate future political activity. Seventh graders are more likely to anticipate participation than are tenth graders.

The four media variables make a substantial contribution to predicting interpersonal discussion ($IR^2 = .17$). Television news viewing is once again the strongest predictor ($\beta = .25$), with newspaper news reading and radio news listening also significant predictors. For anticipated participation, newspaper news reading is most closely related ($\beta = .22$), with the remaining three indices failing to reach significance. Again, addition of the interest variable into the regression equations has a negligible impact on the media coefficients.

CONTINGENT ANALYSES

Contingent correlations were computed between the broadcast and print exposure and criterion variables at high versus low levels of grade, scholarly orientation, and political interest. The expectation that the contingent variables would significantly mediate the relationship between print media exposure and the criterion variables was not supported. Only one such contingent relationship emerges, and in the direction opposite to that predicted. The correlation between print media exposure and anticipated participation is stronger for students at a low level of scholarship than for those at a high level. This contingent relationship is also evident for broadcast media exposure.

Discussion

The present study sought to examine the impact of four types of media exposure—television news, radio news, newspaper news, and newsmagazines—on two types of political knowledge and two forms of political behavior. In addition, the relationships between media exposure and the criterion variables were examined for subgroups of young persons at various levels of age, scholarly orientation, and political interest.

The results of both correlation and regression analyses support the expectation that broadcast news exposure is more strongly related to overall political knowledge than is print news exposure. Also as predicted, broadcast exposure is a better predictor of current events knowledge than of fundamental knowledge. Contrary to expectations, however, print news exposure predicts the two types of knowledge equally well. The predictions regarding the behavioral variables were supported by the data. Broadcast news exposure is more closely related to interpersonal discussion than to anticipated participation, and print news exposure is more strongly related to participation than to discussion.

Only the prediction that print news exposure would be more strongly related to fundamental than to current events knowledge failed to be supported by the data. This failure may be attributed to some of the same characteristics of print news media that led to the original prediction. These characteristics, when interpreted in light of youngsters' limited information-processing abilities, become disadvantages rather than advantages of print for youngsters' gaining fundamental political knowledge. For example, the *in-depth coverage* of print news may result in information overload for the less developed mind. Similarly, the *complexity* of print news stories may render them less comprehensible. The advantage of *self-paced processing* of print applies only to motivated people; the less interested youth may simply fail to capitalize on it.

If processing limitations inhibit youngsters' learning from the print media, then variables indicative of level of processing may mediate fundamental knowledge gain from print news media. Thus older students, brighter students, and more interested students might more efficiently extract fundamental knowledge from the broadcast media than the younger, less bright, and less interested students. The contingent relationships for fundamental knowledge (not shown) do not reveal any consistent pattern in support of this argument, however. But age, scholarship, and political interest are only indirect indicators of processing. More direct measurement of processing suggests that "intense" processing facilitates learning from television news (Garramone, 1983), but whether it facilitates learning sufficiently to allow the extraction of the more fundamental political knowledge should be tested.

The overall pattern of findings across subgroups provides little evidence of unequal effects of exposure; the news media do not appear to produce knowledge gaps (or behavior gaps) between those adolescents who are brighter or politically interested and those who are less scholarly or uninterested. If anything, the media serve a compensatory function in some cases by more strongly stimulating the political orientations of exposed students scoring low on these background variables.

References

Atkin, Charles K.
1978 "Broadcast news programming and the child audience." Journal of Broadcasting 22:47–61.
1981 "Communication and political socialization." Pp. 299–328 in Dan Nimmo and Keith Sanders (eds.), Political Communication Handbook. Beverly Hills, CA: Sage.
Atkin, Charles K., and Walter Gantz
1978 "Television news and the child audience." Public Opinion Quarterly 42:183–98.
Chaffee, Steven H., Scott Ward, and Leonard Tipton
1970 "Mass communication and political socialization." Journalism Quarterly 47:647–59.
Chaffee, Steven H., and Albert R. Tims
1982 "The development of news media use in adolescence: implications for political cognitions." In Michael Burgoon (ed.), Communication Yearbook 6. Beverly Hills, CA: Sage.
Conway, M. Margaret, A. Jay Stevens, and Robert Smith
1975 "The relation between media use and children's civic awareness." Journalism Quarterly 52:531–38.
Drew, Dan, and Byron Reeves
1980a "Children and television news." Journalism Quarterly 57:45–54.
1980b "Learning from a television news story." Communication Research 7:121–35.
Garramone, Gina M.
1983 "Television news and political socialization." In Robert Bostrom (ed.), Communication Yearbook 7. Beverly Hills, CA: Sage.
Gollin, Albert, and Thelma Anderson
1980 America's Children and the Mass Media. New York: Newspaper Advertising Bureau.
Nunnally, Jum
1967 Psychometric Theory. New York: McGraw-Hill.
Sheinkopf, Kenneth
1973 "Family communication patterns and anticipatory socialization." Journalism Quarterly 50:24–30.

Not Forbidding Isn't Allowing: The Cognitive Basis of the Forbid-Allow Asymmetry

HANS-J. HIPPLER AND NORBERT SCHWARZ

SURVEY RESEARCH has repeatedly demonstrated that minor changes in the wording of a question can have a major impact on the obtained responses (see Kalton and Schuman, 1980, and Schuman and Presser, 1981 for reviews). A wording effect that has received much attention in more than four decades of research is the forbid-allow contrast identified by Rugg (1941). He reported that Americans were more likely to support freedom of speech when the appropriate question was worded, "Do you think the United States should forbid public speeches against democracy" (yes/no?) rather than, "Do you think the United States should allow public speeches against democracy" (yes/no?). Specifically, only 54 percent of the respondents who were asked

---

Abstract Previous research demonstrated that respondents are more likely to endorse the idea that something should "not be allowed" (or "not be forbidden") than to endorse the idea that it should be "forbidden" (or "allowed"), even though these expressions seem logically equivalent. The hypothesis is advanced that this asymmetry is due to the response behavior of indifferent respondents who neither endorse that something should be forbidden nor that it should be allowed, resulting in higher endorsements of the negative form of both question wordings. Data consistent with that explanation are presented and the cognitive mechanism underlying the response behavior of indifferents is discussed.

Hans-J. Hippler is Project Director, Zentrum für Umfragen, Methoden und Analysen (ZUMA), Mannheim, and Norbert Schwarz is Assistant Professor of Psychology, Universität Heidelberg, Federal Republic of Germany. This research was partially supported by the Zentrum für Umfragen, Methoden und Analysen and grant Schw 278/2 from the Deutsche Forschungsgemeinschaft to Norbert Schwarz and Fritz Strack. The authors gratefully acknowledge the stimulating discussions with Charles F. Cannell, Howard Schuman, Seymour Sudman, and Fritz Strack. Correspondence should be addressed to H. J. Hippler, ZUMA e.V., Postfach 5969, D-6800 Mannheim 1, West Germany.

Public Opinion Quarterly Vol. 50:87–96 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-87/$2.50

the forbid form of the question felt that those speeches should be "forbidden," but 75 percent of the respondents who were asked the allow form of the question felt that they should "not be allowed." Similarly, only 25 percent of the respondents felt that these speeches should be "allowed," but 46 percent responded they should "not be forbidden."

Since this original study, the effect has been successfully replicated over a wide range of issues, and Schuman and Presser (1981:296), in reviewing this work, noted that it is "the largest wording effect we have discovered," producing differences of up to 30 percentage points. Three factors that possibly contribute to the effect have been suggested. First, and most notably, "forbid" and "allow" carry different connotations: "the former sounds harsher and may therefore be more difficult to endorse, whereas the latter in some context might seem to encourage a deviant behavior and therefore may invite opposition" (Schuman and Presser, 1981:280). Second, the terms create different grammatical structures, and finally, the effect may be limited to abstract issues such as speeches against democracy. Unfortunately, none of these explanations received unequivocal support in studies reported by Schuman and Presser, and they conclude that they "are hard put to see any obvious source for the effect" (1981:296).

However, it seems to us that these authors discarded the connotation explanation somewhat prematurely on the basis of a potentially inadequate test. Specifically, Schuman and Presser (1981:281) found large and reliable wording effects for abstract items (e.g., speeches against democracy) but small and unreliable effects for more concrete items (e.g., the showing of X-rated movies, or cigarette advertisement on TV). They concluded from this pattern: "If forbid and not allow (or not forbid and allow) convey somewhat different meanings, the effect of the verbal difference taken alone is not large. For the alternative formulations to have much impact, it would appear that they have to be joined to abstract issues like free speech or communism" (1981:282). Interestingly, Schuman and Presser also reported higher don't know levels for their abstract than for their concrete items, and they concluded that "the higher the don't know level the larger the forbid-allow effect" (p. 283). To the extent that high don't know levels indicate a high level of indifference in the public opinion climate (Noelle-Neumann, 1980), this finding suggests a variant of the connotation explanation that might well account for the data reviewed by Schuman and Presser (1981) in the manner outlined below.

Specifically, we want to suggest that the forbid-allow asymmetry is primarily due to the response behavior of indifferent (or ambivalent) respondents. We assume that respondents who hold a strong attitude toward the issue under study will give consistent responses to both forms of the question. That is, respondents opposed to the issue will endorse the forbid statement as well as the not allow statement, while those in favor of the issue will endorse both the allow and not forbid statements. However, indifferent respondents may respond "no" to *both* questions. Specifically, when asked whether something should be forbidden, indifferents may say "no" because they are not opposed to the issue but only indifferent. However, when asked whether it should be allowed they may respond "no" for the same reason—even though a "no" response seems to imply different opinions depending on which question it is given to. Two factors may contribute to this phenomenon, namely, the connotations of "forbid" and "allow," and a cognitive bias to focus on positive instances of a behavior rather than to consider the implications of its absence.

With regard to the connotations of "forbid" and "allow" it has already been noted that to "forbid" is a rather harsh statement while "allowing" something may be construed as inviting it, rather than tolerating it. That is, to *forbid* something implies stronger opposition than *not to allow* something, because forbidding refers to an active act of opposition against the issue under consideration while not allowing something only refers to abstaining from active support. Similarly, *allowing* something refers to active support while *not forbidding* something refers to abstaining from active opposition.

Thus, respondents who do not hold a strong attitude may be unlikely to endorse *either* the forbid or the allow statement, responding "no" to both question forms. In doing so they seem to endorse the opposite position. However, they may miss this in the interview situation because they focus on the behavior they are asked about. In this regard, research in the area of social cognition demonstrated that individuals pay more attention to positive instances of a behavior than to the absence of a behavior (cf. Nisbett and Ross, 1980 for a review). For example, they feel that what a person does do provides more information about the person than what the person does not do (e.g., Fazio et al., 1982). This "tendency of both animals and humans to exhibit greater difficulty in the processing of nonoccurrences than occurrences" (Fazio et al., 1982:404) is usually referred to as the *feature-positive effect.* Similarly, when asked if something should be done (allowed or forbidden), individuals may focus on the implications of the behavior under consideration, and they may *not* consider the implications of the *absence* of this behavior. Such a bias could contribute to the forbid-allow asymmetries by clouding the idea that not allowing or not forbidding something indeed amounts to forbidding or allowing it, respectively, thus resulting in the feeling that one has only said it shouldn't be allowed without ever commenting on whether it should be forbidden.

Note that this explanation predicts that the asymmetries are primarily due to the behavior of indifferent or ambivalent respondents and should not be obtained for respondents holding a clear pro or con attitude. In line with this hypothesis, the forbid-allow effect was found to be more pronounced the more abstract the issue and the lower the education of the respondents, both factors that may affect the number of indifferent respondents, which is also partially reflected in increasing don't know levels.

In this article we will report evidence that to forbid and to allow are stronger statements than not to allow or not to forbid (Experiment 3), and that the forbid-allow asymmetry is limited to indifferent respondents and is not obtained for respondents holding a strong attitude (Experiment 4). Because our studies were conducted in West Germany, however, we will first demonstrate that the basic forbid-allow effect is cross-culturally replicable (Experiments 1 and 2).

Experiments 1 and 2: The Forbid-Allow Effect in German Samples

To test the cross-cultural stability of the forbid-allow effect, two split-ballot experiments were conducted. The first experiment was carried out in July 1982 as part of the second wave of a small face-to-face panel survey with a random sample of adults (more than 18 years of age) living in an industrial city in West Germany and selected from registries (response rate first wave: 63 percent; second wave: 89 percent). The second experiment was carried out in April 1983 as part of a larger face-to-face survey (quota sample) with 146 German adults. Quotas were based on an intersection of age, sex, and years of education. In the first experiment, 48 of 88 respondents were asked if peep shows should be allowed (yes or no), and 40 were asked if they should be forbidden (yes or no). To provide a stimulus replication, 146 respondents in the second survey were asked if the showing of X-rated movies in public cinemas should be allowed ($N = 72$) or forbidden ($N = 74$). Both peep shows and X-rated movies are legal in West Germany.

Both studies replicated the forbid-allow asymmetry, demonstrating its cross-cultural stability. Specifically, 27.7 percent and 26 percent of the respondents, respectively, endorsed the statement that peep shows and the showing of X-rated movies should be forbidden, while 50 percent and 40 percent, respectively, reported they should not be allowed, χ^2 (1) = 3.5 and 2.9, respectively, $p < .06$ and .10 (three and two subjects, respectively, provided a don't know response). Thus, differences of 23.3 and 14 percentage points were obtained as a function of question wording, using two issues that qualify as concrete rather than abstract in Schuman and Presser's terminology. As in Schuman and Presser's (1981:282) data for X-rated movies, the effects were only marginally reliable. However, a combined analysis of both studies (following a suggestion by Rosenthal, 1978) indicates the statistical reliability of the general forbid-allow asymmetry in both German samples, $z = 2.49$, $p < .01$, one-tailed.

Experiment 3: Forbidding and Allowing Are Stronger Statements Than Not Allowing or Not Forbidding

To test the hypothesis that to allow something or to forbid something are stronger positions to take than not to forbid or not to allow it, ratings of statements containing these expressions were obtained.

METHOD

As part of a self-administered questionnaire 54 adults (more than 18 years of age) of a quota sample (quotas were based on an intersection of age, sex, and years of education) read a sentence stating that "Mr. Mueller feels the legislator should forbid (not allow, allow, not forbid) peep shows," and subsequently rated Mr. Mueller's opinion on an 11-point bipolar scale labeled "extremely in favor of peep shows" vs. "extremely opposed to peep shows." Each respondent read and rated only one statement.

RESULTS

As expected, "allowing" peep shows was rated as representing a more favorable attitude, $M = 2.1$, than "not forbidding" peep shows, $M = 3.1$, $t(50) = 2.14$, $p < .04$. Similarly, "forbidding" peep shows was rated as indicating stronger opposition, $M = 10.5$, than "not allowing" them, $M = 9.2$, $t(50) = 2.69$, $p < .01$, demonstrating that the terms are perceived as differing in extremity even though they may be logically equivalent in their consequences. The differences are small, however, and there is no doubt that "not forbidding" is seen as a statement in favor of the issue while "not allowing" is seen as a statement in opposition to the issue.

Experiment 4:
The Forbid-Allow Effect Is Limited to Indifferent Respondents

Central to the present argument is the assumption that indifferent respondents may respond negatively to both a question asking them whether something should be allowed and a question asking them whether it should be forbidden. Moreover, it is assumed that this response behavior accounts for the higher overall endorsement of the negative form of both questions. To test this hypothesis, respondents in favor of, opposed to, or indifferent toward spreading salt on the roads for safety in winter—a popular topic in the current ecological discussion in West Germany—were asked whether using salt to melt ice and snow should be allowed (yes or no) or forbidden (yes or no), resulting in a 3 × 2 factorial design with attitude (pro, con, indifferent) and question wording (forbid, allow) as between subject factors.

It was expected (a) that more respondents overall would endorse the "not allow" than the "forbid" and the "not forbid" than the "allow" statements, and (b) that this would be primarily due to the responses of indifferent subjects, and would not be obtained in the responses of subjects holding a strong attitude.

METHOD

A random sample (selected from registries) of 720 adults (more than 18 years of age) living in an industrial city in the Federal Republic of Germany participated in a mail survey (response rate 78 percent) concerned with environmental issues. Early in the questionnaire (question number 10) respondents reported their attitudes toward 12 environmental issues, including the use of salt for road safety in winter. Specifically, the question read, "Salt should no longer be used to melt ice and snow on the roads." Respondents were asked to check one of three response alternatives ("agree," "don't agree," "indifferent/don't care").

Later in the questionnaire (question number 34) half of the respondents were asked whether the use of salt to melt ice should be allowed (yes or no), while the remaining respondents were asked whether it should be forbidden (yes or no). The question read: "The use of salt during the winter is dangerous to plants and trees near the road. Do you think that one should generally allow (forbid) the use of salt (yes/no)?" A "don't know" option was not provided with this question.

RESULTS

Table 1 presents the responses of the total sample as a function of question wording. As expected, more respondents endorsed the "not allow" than the "forbid" statement and the "not forbid" than the "allow" statement, $\chi^2(1) = 23.2, p < .001$. Thus, the data replicate the forbid-allow asymmetry resulting in a wording effect of 16.5 percentage points.

Separate analyses of the responses provided by respondents who

Table 1. The Forbid–Allow Asymmetry as a Function of Respondents' Attitude

Forbid Form		*Allow Form*	
Do you think that one should generally forbid the use of salt?		Do you think that one should generally allow the use of salt?	
Yes (forbid)	62.3% (215)	No (not allow)	78.8% (283)
No (not forbid)	37.7% (130)	Yes (allow)	21.2% (76)
	100 (345)		100 (359)
Response × form $\chi^2 = 23.2$, df = 1, $p < .001$, Don't know: $N = 16$			
Attitude Opposed to the Use of Salt			
Yes (forbid)	86.4 (186)	No (not allow)	90.7 (215)
No (not forbid)	13.6 (29)	Yes (allow)	9.3 (22)
	100.0 (215)		100.0 (237)
Response × form $\chi^2 = 2.0$, df = 1, n.s.			
Attitude in Favor of the Use of Salt			
Yes (forbid)	25.8 (17)	No (not allow)	38.7 (24)
No (not forbid)	74.2 (49)	Yes (allow)	61.3 (38)
	100.0 (66)		100.0 (62)
Response × form $\chi^2 = 2.5$, df = 1, n.s..			
Attitude Indifferent to the Use of Salt			
Yes (forbid)	18.8 (12)	No (not allow)	73.3 (44)
No (not forbid)	81.2 (52)	Yes (allow)	26.7 (16)
	100.00 (64)		100.0 (60)
Response × form $\chi^2 = 37.3$, df = 1, $p < .001$			

had reported a pro, con, or indifferent attitude toward the use of castor salt are presented in the lower part of Table 1. A total of 89 percent of the respondents who reported an opposed attitude toward castor salt also wanted the use of castor salt to be discontinued. As expected, their responses were *not* significantly affected by question wording $\chi^2(1) = 2.0$, n.s., and only a small effect of 4.3 percentage points was obtained. Similarly, no significant wording effect was obtained for respondents in favor of the use of castor salt, $\chi^2(1) = 2.5$, n.s., though the difference of 13.1 percentage points was somewhat larger than for respondents holding an opposed attitude. Moreover, 32 percent of the respondents who had reported a favorable attitude on the earlier question nevertheless wanted the use of castor salt to be discontinued. This might be due to the information about its negative effects on plants and trees provided in the question text, which was not provided in the earlier attitude question.

Finally, a large and highly significant wording effect of 54.5 percentage points, $\chi^2(1) = 37.3$, $p < .001$ was obtained for respondents who had reported an indifferent attitude. Specifically, these individuals were likely to respond no to *both* the "allow" (81.2 percent) and the "forbid" (73.3 percent) form of the question, as suggested by our theoretical analysis. Thus, the present data suport the hypothesis that the forbid-allow asymmetry is primarily due to the response behavior of indifferent respondents who will neither endorse that something should be allowed nor that it should be forbidden, because they want to express neither support nor opposition.

While these results are in line with our theoretical analysis, a methodological issue requires additional consideration. To identify indifferents, respondents' attitude toward the use of salt was assessed independently of and prior to the experimental question. Though the questions were separated by 23 unrelated ones, this raises the possibility of consistency effects. Most important, respondents who first reported being indifferent about the issue may be particularly unlikely to endorse a "strong," i.e., affirmative, option later on. Thus, a consistency effect might increase the size of the forbid-allow asymmetry for indifferents. While the present data do not allow us to rule out this possibility, it seems unlikely that a consistency effect alone could account for the pattern of results obtained. Moreover, the fact that 32 percent and 11.3 percent of the respondents who initially reported a favorable or an opposed attitude, respectively, gave an opposite response later on suggests that consistency pressures could not have been very strong.

General Discussion

We have reported here evidence for the stability of the forbid-not allow and allow-not forbid asymmetries across cultures and languages by replicating the effect in Germany (Experiments 1 and 2). Moreover, we attempted to demonstrate that these asymmetries are due to the behavior of respondents who are relatively indifferent about the issue under consideration (Experiment 4). Specifically, we suggested that respondents holding a strong attitude in favor of or in opposition to the issue will give consistent responses independent of question wording. Indifferent respondents, on the other hand, may refuse to endorse *either* that something should be allowed or that it should be forbidden because they are neither strongly in favor of nor strongly opposed to the issue. Therefore, they will be lumped together with respondents holding a strong attitude, thus resulting in the asymmetries under study, with more respondents endorsing the negative than the affirmative form of each question.

Two factors may contribute to the response behavior of indifferents. Most important, respondents are likely to focus on the implications of *doing* what they are asked about (that is, allowing or forbidding something) rather than not doing it, thus exhibiting a well-documented bias to focus on positive instances of a behavior (cf. Fazio, et al., 1982; Nisbett and Ross, 1980). Therefore, respondents may miss the implications of their negative response, feeling that they only gave their opinion about allowing (forbidding) something without ever commenting on the opposite. Moreover, the negative wordings are weaker statements than the affirmative wordings (Experiment 3) and may thus be easier to endorse for indifferents.

The present explanation of the forbid-allow asymmetries implies that the effect should be more pronounced in aggregated data the larger the number of indifferents. Therefore, the effect should be the larger the more abstract the issue, the less salient the issue, and the lower the education of the respondents. Indeed, the data reviewed by Schuman and Presser (1981) provide evidence in support of each of these variables. In addition, the present explanation suggests that the forbid-allow asymmetries may vanish if indifferent respondents are provided the possibility of expressing their indifference. To the best of our knowledge such a study has not been conducted and the "don't know" option provided in some of Schuman and Presser's (1981) studies seems inappropriate to attain this goal. If respondents do indeed focus on the behavior described in the question, that is, to forbid or to allow, indifferents *do know* that they do not want to forbid or to allow something. Thus, responding "don't know" would not accommodate their need to express indifference. "Don't care" or "leave things as they are" categories might be more appropriate to identify indifferents, and to avoid their being lumped together with respondents holding an attitude they do not share.

References

Fazio, R. H., S. J. Sherman, and P. M. Herr
1982 "The feature positive effect in the self-perception process: does not doing matter as much as doing?" Journal of Personality and Social Psychology 42:404–11.

Kalton, G., and Howard Schuman
1980 The effect of the question on survey responses: a review. Proceedings of the Section on Survey Research Methods, Washington, D.C.: American Statistical Association.

Nisbett, R, and Lee Ross
1980 Human Inference: Strategies and Shortcomings of Social Judgment. Englewood Cliffs, NJ: Prentice-Hall.

Noelle-Neumann, E.
1980 Die Schweigespirale. München: Piper.

Rosenthal, R.
1978 "Combining results of independent studies." Psychological Bulletin 85:185–93.

Rugg, D.
1941 "Experiments in wording questions: II." Public Opinion Quarterly 5:91–92.
Schuman, H., and S. Presser
1981 Questions and Answers in Attitude Surveys. New York: Academic Press.

Appendix A

German text of the questions used in the reported experiments

EXPERIMENT 1

Meinen Sie, der Gesetzgeber sollte die sogenannten Peep-Shows, also das bezahlte Sichzurschaustellen nackter Frauen in aufreizenden Posen, generell verbieten (erlauben)? (Ja/Nein)

EXPERIMENT 2

Meinen Sie, der Gesetzgeber sollte die Vorführung von pornographischen Filmen in öffentlichen Kinos generell verbieten (erlauben)? (Ja/Nein)

EXPERIMENT 3

Wir hätten gerne von Ihnen eine Beurteilung der folgenden Aussage. Herr Müller meint: "Der Gesetzgeber sollte die sogenannten Peep-Shows, also das bezahlte Sichzurschaustellen nackter Frauen in aufreizenden Posen, generell verbieten (erlauben/nicht erlauben/nicht verbieten)."

Wie sehr ist Ihrer Ansicht nach Herr Müller für bzw. gegen Peep-Shows? Machen Sie bitte ein Kreuz auf der folgenden Skala.

+5	+4	+3	+2	+1	0	−1	−2	−3	−4	−5

extrem *für*
Peep-Shows

extrem *gegen*
Peep-Shows

EXPERIMENT 4

Q.10. Wir hätten gerne von Ihnen gewuβt, wie Sie zu den folgenden Aussagen stehen. Stimmen Sie der jeweiligen Aussage zu, ist sie Ihnen egal oder lehnen Sie sie ab? Streusalz sollte bei Glatteis nicht mehr verwendet werden (ja, stimme zu; ist mir egal/weiβ nicht; nein, lehne ab)

Q.34. Durch das Streuen von Salz im Winter sind Pflanzen und Bäume in Straβennähe gefährdet. Sollte man das Salzstreuen generell erlauben (verbieten)? (Ja/Nein)

Interviewers' Voices and Refusal Rates in Telephone Surveys

LOIS OKSENBERG, LERITA COLEMAN
AND CHARLES F. CANNELL

As SURVEY RESEARCH uses telephone interviews more extensively, new methodological problems become apparent. One problem is response rates, which are generally lower than for face-to-face interviews. The reasons for the lower response rates are not known. Refusals in face-to-face interviews are often attributed not only to the content but to the respondent's perception of the interviewer during initial introductory phrases and impressions based on his or her appearance. In phone interviews these latter cues are absent, and auditory stimuli are the sole basis for evaluation. This implies that characteristics of voice and speech patterns provide major bases for first impressions, and may relate to acceptance or rejection of the interview.

In listening to telephone interviewers at work one is impressed with their unique speech characteristics. Further, interviewers are characterized by high or low response rates, and they tend to maintain their

---

Abstract Interviewers' voice characteristics are related to refusal rates in telephone interviews. Judges reliably rated vocal and personal characteristics of telephone interviewers from recordings of the first one-half minute of interview introductions. Interviewers with low refusal rates were judged as having comparatively higher pitched voices, greater ranges of variation in pitch, greater loudness, faster rates of speaking, and clearer and more distinct pronunciation. They also were judged as more competent and as having a more positive approach to the respondent and the interview.

The authors are with the Institute for Social Research of The University of Michigan: Lois Oksenberg is a Research Investigator and Charles F. Cannell is Program Director Emeritus in the Survey Research Center; Lerita Coleman is an Assistant Research Scientist in the Research Center for Group Dynamics. The authors are grateful to Maryellen McSweeney for her statistical advice. This material is based upon work supported by the National Science Foundation under Grant No. SES-8112801.

Public Opinion Quarterly Vol. 50:97–111 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-97/$2.50

rate over time. For example, Survey Research Center (The University of Michigan) telephone interviewers working on a continuing monthly survey showed considerable consistency in refusal rates over time, with rates ranging from 6 to 42 percent.

These observations led us to attempt a systematic analysis of interviewers' voice characteristics to see whether we could identify factors correlated with refusal rates. If factors could be identified, we might be able to select interviewers with positive voices or train interviewers to speak more acceptably.

Of the nonresponses in telephone surveys, one-half to two-thirds are refusals to be interviewed. The great majority of refusals occur within the opening minutes of the interview, prior to the first substantive question. In a study at the Survey Research Center (SRC), for example, about a third of the refusals occurred in the opening seconds, another third occurred during the body of the introduction, while the remainder occurred at the point of listing household members prior to the first question. Respondents appear to form an attitude toward the interview early and accept or refuse to participate accordingly.

Efforts to improve response rates have focused on the content of introductions, but have met with little success (Dillman et al., 1976; O'Neil et al., 1979). Differences in refusal rates among interviewers appear to reflect factors unrelated to content. Listening to interview introductions reinforced the idea that interviewers' vocal characteristics are key to the impressions potential respondents form of them and whether they would consent to and would enjoy an extended interview with them.

Initially a person probably will have a somewhat negative attitude toward a telephone call from a stranger. The caller is likely to be selling something, soliciting contributions, making a political appeal, or imposing on the recipient in some other way. As conversation progresses, the recipient may change his or her attitude toward the call. If the purpose of the call is attractive, he or she may become less negatively or even positively disposed. If the purpose is unattractive or unclear, the recipient will become more negative. The characteristics exhibited by the caller will also affect the recipient's attitude. There will be less interest in engaging in extended conversation with an unappealing caller.

The telephone interviewer shares these barriers to communication. The presentation must be sufficiently positive to overcome initial indifference or negative feelings. In a face-to-face interview there are a number of cues to help establish a positive attitude: The interviewer's vocal characteristics, physical appearance, dress, demeanor, interviewing paraphernalia and, perhaps, motor vehicle are all evident. In contrast, a telephone interviewer must create a sufficiently positive impression with only vocal cues. If vocal characteristics lead the respondent to perceive the interviewer as unappealing, cooperation will be less likely.

It is evident from the history of research on paralanguage and language attitudes that listeners rely heavily on vocal characteristics to form impressions and attitudes about speakers (Kramer, 1963; Giles and Powesland, 1975; Giles et al., 1980). Judges tend to agree in rating speakers' personal attributes from their voices (Addington, 1968; Allport and Cantril, 1934; Kramer, 1964). While rater agreement appears to reflect vocal stereotypes rather than accurate judgments, it nonetheless indicates that people hold definite ideas about the correlation of vocal attributes with personality.

Judgments about demographic characteristics are based on vocal cues. Sex, age, social status, and race are accurately identified from the voice (Smith, 1979; Pear, 1931; Harms, 1961). These demographic characteristics appear to influence judgments of personal characteristics. Low ascribed status, for example, is associated with less credibility (Harms, 1961; Putnam and O'Hearn, 1955; Ellis, 1967). Recent research has noted that, in general, female speakers are perceived to be more socially attractive and competent than male speakers (Smith, 1979).

A number of studies, primarily using male speakers, have attempted to isolate and identify specific vocal cues involved in evaluation. These studies have utilized a number of vocal features including pitch, rate of speech, amplitude (loudness), intonation (variation in pitch), pauses, and fluency. Increases in pitch, for example, produce judgments of less competence and less benevolence, while faster speech elicits ratings of competence (Brown et al., 1973). Greater variability in pitch, more speech volume, increased speech rate and less halting speech have been linked to persuasiveness (Mehrabian and Williams, 1969; Packwood, 1974). Rapid speech rates elicit ratings of more credibility and enhanced persuasiveness (Pearce and Conklin, 1971). Faster speakers are rated as more intelligent, knowledgeable, and objective than slower speakers (Miller et al., 1976). Perceived persuasiveness may not result from a fast rate of speech per se but from a smooth, nonhesitant speaking style. Large numbers of pauses in speech produce negative ratings of speakers (e.g., as angry, anxious, contemptuous, and bored [Lalljee, 1971]). Other highly desirable traits, including credibility, are associated with nonhesitant speakers (Lay and Burron, 1968).

The effect of vocal characteristics on personal attributions may depend on the sex of the speaker. Addington (1968) found that greater variation in pitch was interpreted somewhat differently for male and female speakers. To further complicate matters, Apple et al. (1979) found amplitude, pitch, and rate of speaking to be intercorrelated in normal speech, making it difficult to link particular characteristics to perceptions of the speaker.

The literature on vocal cues prompted a number of specific questions about the effects of voices of telephone interviewers. What cues do respondents use to judge interviewers? Is there agreement on the characteristics of voices and on the personal attributes signalled by them, or are these perceptions so idiosyncratic that no general patterns emerge? If there is agreement on the voice and personal attributes signalled by voices, are these perceptions related to interviewer refusal rates?

Three objectives for this study were derived from these questions: (1) Can vocal qualities of telephone interviewers be judged reliably? (2) Do ratings of vocal qualities correlate with perceptions of desirable and undesirable personal attributes? (3) Is there a relation between these characteristics and rate of refusals?

This article reports our first investigations of these issues. In two studies, raters listened to recordings of the first one-half minute of telephone interview introductions and evaluated a number of the interviewers' vocal and personal characteristics. Because of the small number of male interviewers on the SRC staff and the need to restrict the scope of the research, only female interviewers were included. In addition to the rating studies, acoustical analyses of physical properties of interviewer voices were performed and reported by Sharf and Lehman (1984).

Measures of Vocal and Personal Characteristics

Rating scales were developed by the project staff on the basis of previous paralinguistic, personality, person-perception, and attitude change studies (Figures 1 and 2). Our goal was to represent key vocal cues and personal attributes identified in the literature, as well as characteristics that, on the basis of our own evaluations, seemed particularly relevant to telephone interviewing. We included only the number of vocal or personal characteristics that could be rated in a single session. After extensive investigation, 10 scales measuring vocal characteristics and 21 scales measuring personal attributes were selected. With the exception of a five-category scale for the interviewer's social class and a six-category scale for age, all scales were seven-point anchored semantic differential type.

Figure 1. Rating Interviewer's Voice Quality

VOICE NUMBER ○

Please rate the voice on the following scales.
For each scale, circle the number that best describes what you hear.

Average pitch	low	1 2 3 4 5 6 7	high
Variation in pitch (intonation)	small range	1 2 3 4 5 6 7	large range
Loudness	soft	1 2 3 4 5 6 7	loud
Stress or emphasis of words and phrases	done appropriately	1 2 3 4 5 6 7	not done appropriately
Flow of words	smooth and easy	1 2 3 4 5 6 7	hesitant
Pronunciation of words	clear and distinct	1 2 3 4 5 6 7	indistinct or slurred
Spontaneity	sounds natural/ spontaneous	1 2 3 4 5 6 7	sounds memorized or like reading or otherwise "canned"
Rate of speaking	slow	1 2 3 4 5 6 7	fast
Any sort of accent? Please check: ____ NO ____ YES. IF YES:	pleasant	1 2 3 4 5 6 7	unpleasant
Generally, how pleasant was this voice to listen to:	pleasant	1 2 3 4 5 6 7	unpleasant

Recordings

Introductory statements made by experienced SRC interviewers were recorded. The recordings of actual telephone interviews were made through an extension phone in order to reproduce as closely as possible the voice as heard by the respondent. Interviewers used the standardized introduction provided for the survey with little variation. The recorded portion read:

> Hello, my name is ______. I'm calling from The University of Michigan, in Ann Arbor. Here at the university we are currently working on a study for the Survey Research Center. First of all I need to be sure I've dialed the right number. Is this ______ (telephone number)? Since this telephone number has been generated by a computer, I do not know whether this number is for a business or a home. Which is it?

The recordings were approximately 30 seconds long. Interviewers' names and respondents' area codes were deleted from recorded introductions.

Figure 2. Rating Interviewer's Personal Characteristics

Interviewer Number ○

Please rate the interviewer's manner and personal characteristics on the following scales. For each scale, circle the number that best describes your perception of the interviewer.

young	1 2 3 4 5 6 7	old
cheerful	1 2 3 4 5 6 7	uncheerful
good at her job	1 2 3 4 5 6 7	poor at her job
interesting	1 2 3 4 5 6 7	boring
self-assured	1 2 3 4 5 6 7	ill at ease
educated	1 2 3 4 5 6 7	uneducated
trustworthy	1 2 3 4 5 6 7	untrustworthy
friendly	1 2 3 4 5 6 7	unfriendly
intelligent	1 2 3 4 5 6 7	unintelligent
enthusiastic	1 2 3 4 5 6 7	apathetic
impressive	1 2 3 4 5 6 7	unimpressive
interested in her task	1 2 3 4 5 6 7	bored with her task
polite	1 2 3 4 5 6 7	impolite
professional	1 2 3 4 5 6 7	amateurish
kind	1 2 3 4 5 6 7	unkind
mature	1 2 3 4 5 6 7	immature
pleasant	1 2 3 4 5 6 7	unpleasant
good-looking	1 2 3 4 5 6 7	plain

What do you think the interviewer's social class is? (Please check)

____ working
____ lower-middle
____ middle
____ upper-middle
____ upper

What do you think the interviewer's age is? (Please check)

____ under 20
____ 20–29
____ 30–39
____ 40–49
____ 50–59
____ 60–69

How would you like to be interviewed *by this person?*

like to be interviewed by her	1 2 3 4 5 6 7	dislike to be interviewed by her

The Two Studies

Study 1 was based on recordings of six interviewers; three interviewers with low refusal rates (.03, .11, .14) and three with high refusal rates (.36, .38, .46) over a seven-month period were selected. A seventh interviewer with an intermediate refusal rate was included for practice ratings.

Eighteen students (11 women and 7 men) in a summer course in survey methodology at The University of Michigan served as raters. They ranged in age from the twenties to the fifties. Ratings of the vocal characteristics for each of the six voices were obtained in one session, and ratings of the personal characteristics for the same six voices were obtained in another session the following day. This minimized influence between ratings of the two types of characteristics and avoided too great a burden on raters.

On the first day, after a brief introduction to their task, raters heard the practice voice twice, and then had about one minute for rating. The same procedure was followed for the remaining voices. Voices of the three more successful interviewers and the three less successful interviewers were alternated. The procedure in the second session, in which personal characteristics were rated, followed the same pattern.

A second study (Study 2) used raters recruited from a local church. The raters ranged in age from 18 to over 60 and were primarily working class or middle class and predominantly white men and women (22 females and 18 males). As in the first study, ratings were obtained for six interviewers: three with low refusal rates (.10, .13, and .14) and three with high refusal rates (.24, .29, and .38). Two of the six voices and the practice voice had been included in the first study. Procedures were the same as those in Study 1, except that the two rating sessions were on consecutive Sundays, and the raters received an introduction to the task better geared to a general audience.

Construction of Measures and Reliability

Rating scale reliabilities were obtained for each of the studies. The analysis of variance method with repeated measures giving mean interrater reliabilities as described by Winer (1971) was applied. Reliabilities of the mean ratings assigned by the listeners were generally very high—the great majority in the .80s and .90s in both studies. High correlations among many of the ratings, however, indicated considerable overlap among scales. Ratings for four vocal characteristics were so highly intercorrelated in both studies that they appeared to measure a single factor. They included *stress or emphasis of words and phrases, flow of words, spontaneity,* and *generally, how pleasant was this voice to listen to.* This conclusion was supported by principal axes factor analyses of the ratings of vocal characteristics undertaken separately for the two studies. Since the object was to investigate relations among ratings made by individual listeners, each listener in a study contributed six sets of ratings to the analysis, one set for each voice. Ratings of the four characteristics were combined in one measure referred to as *speaking skill,* scored as the mean of the four ratings. The remaining vocal characteristics showed lower levels of correlation and were kept separate. The scale measuring *pleasantness of accent* was not analyzed because of the small number of times interviewers were rated as having accents.

Personal characteristic ratings showed generally high levels of correlation in both studies as well as some grouping. For all personal characteristics (except *young, social class,* and *age category*) one anchor of the scale was positively valued, the other negatively valued. To explore how best to summarize findings concerning these evaluative scales, we performed principal axes factor analyses including only the 15 evaluative scales that were reliable in both studies.[1] The *polite, kind,* and *trustworthy* scales were not included in these or subsequent analyses. They were quite unreliable in Study 1 since all six voices were rated uniformly high. For each study one factor including all 15 scales provided an economical description of their interrelations. (A single factor accounted for 66 percent of the variance in Study 1, 57 percent in Study 2.) Accordingly, a *general attractiveness* measure was constructed and scored as the mean of ratings on the 15 scales.

Correlations among the 15 scales, however, showed some clustering into two groups: one related to competence of the interviewer, the other to her approach to the respondent and the interview. The theoretical interest of these groups of scales led us to perform additional factor analyses to identify relatively pure measures of the underlying concepts. For each study we extracted two factors and rotated them using an oblique solution. Scales that had very high loadings on the first factor and much lower loadings on the second factor were grouped together, while scales with very high loadings on the second factor and much lower loadings on the first factor likewise were grouped. Groups formed for each of the studies were then compared. Scales grouped together in both studies were used to measure the relevant concepts. The first concept measured was *competence* including scales for *educated, intelligent,* and *professional.* The second concept measured was *positive approach* including scales for *cheerful, friendly, enthusiastic, interested in her task,* and *pleasant.* Both measures were scored as the mean of the ratings on the respective scales.

Table 1 gives the mean inter-rater reliabilities for the measures. The *young* scale was not analyzed because it yielded results indistinguishable from *age category.* The extremely high levels of reliability show that listeners reliably rated vocal and personal characteristics of telephone interviewers from recordings of the first 30 seconds or so of their introductions.[2]

---

[1] To control for differences in listeners' use of the scales, the six ratings for each characteristic contributed by a given listener were standardized by subtracting each rating from the mean of the six.

[2] Levels of agreement between any two raters are lower than the reliabilities reported here. For example, a mean inter-rater reliability of .91 in Study 1 corresponds to an estimated mean correlation between ratings of all pairs of listeners of .35 in Study 1 and of .20 in Study 2 (Rosenthal, 1973). Individuals particularly skilled in making the judgments must be identified or trained in order to obtain reliable ratings from only one or a very few raters.

Table 1. Mean Inter-Rater Reliabilities of Interviewer Characteristic Measures in Two Studies

Scale	*Study 1*[a]	*Study 2*[b]
Vocal characteristics		
Average pitch	.95	.98
Variation in pitch	.88	.89
Rate of speaking	.90	.88
Loudness	.86	.93
Pronunciation	.88	.86
Speaking skill[c]	.92	.89
Personal characteristics		
General attractiveness[d]	.91	.89
Positive approach[e]	.90	.90
Competence[f]	.89	.86
Social class	.83	.78
Age category	.95	.99

[a] Figures are based on 18 raters (except that for *pronunciation*, which is based on 17 raters).

[b] Figures are based on 40 raters (except that for *loudness*, which is based on 36 raters, and those for *variation in pitch* and *rate of speaking*, which are based on 39 raters).

[c] Includes *stress and emphasis on words and phrases, flow of words, spontaneity*, and *generally, how pleasant was this voice to listen to.*

[d] Includes 15 personal characteristics: those in the *positive approach* and *competence* measures, plus *good at her job, interesting, self-assured, impressive, mature, good looking*, and *like to be interviewed by her.*

[e] Includes *cheerful, friendly, enthusiastic, interested in her task*, and *pleasant.*

[f] Includes *educated, intelligent*, and *professional.*

Relation of Rated Characteristics to Refusal Rates

Tables 2 and 3 give the mean ratings for the three interviewers with low refusal rates and for the three interviewers with high refusal rates in each of the two studies. The tables also give results of comparisons of the mean ratings for the two groups of interviewers, one comparison for each measure. The comparisons were tested within the context of the respective analyses of variance used to determine reliabilities. The tables also give η^2, which is the proportion of variation in the ratings attributable to the difference between the two groups of interviewers. Since the significance levels for η^2 are the same as those of the F for the associated comparison, they are noted only for the F values.

The η^2 for all vocal measures differed significantly from zero in both studies, and several were quite large (Table 2). In Study 1 the two elements of pitch (*average pitch* and *variation in pitch*) were strongly related to interviewer refusal rates, with η^2 of .30 and .19, respectively. They were followed by *loudness, pronunciation, rate of speaking*, and *speaking skill*. The η^2 in Study 2 were somewhat lower. *Loudness* had the highest value (.15), followed by *average pitch, variation in pitch,*

Table 2. Mean Ratings of Interviewers Grouped by Refusal Rates, *F*s, η^2, and Direction of Differences for Vocal Characteristics in Two Studies

	Mean Rating: Interviewer Group				
Characteristic	*Low Refusal Rate*	*High Refusal Rate*	*F**	η^2	*Characterization of Interviewers with Low Refusal Rates*
Average pitch	4.76	3.24	64.19[a]	.30	Higher
	4.63	*3.66*	*58.46*[b]	*.11*	pitch
Variation	4.61	3.15	29.83	.19	Greater
in pitch	*4.62*	*3.90*	*24.74*	*.08*	variation
Loudness	4.65	3.76	28.56	.17	Louder
	4.80	*4.06*	*59.90*	*.15*	
Rate of speaking	4.20	3.56	14.93	.07	Faster
	4.01	*3.39*	*24.35*	*.08*	
Pronunciation	5.19	4.23	19.19	.11	Clearer and
	5.69	*4.98*	*21.82*	*.06*	more distinct
Speaking skill	4.22	3.67	7.91	.04	More
	4.62	*4.25*	*6.60*[c]	*.02*	skilled

NOTE: Measures were scored as indicated in Figure 1, except for *pronunciation* and *speaking skill*. Scoring was reversed for these two measures so that high scores indicated clear and distinct pronunciation and good speaking skill. Italicized figures are for Study 2. Other figures are for Study 1.

* Except as noted $p < .01$

[a] For five characteristics in Study 1 $df = 1{,}85$. For one characteristic $df = 1{,}80$.

[b] For three characteristics in Study 2 $df = 1{,}195$. For three others $df = 1{,}190$ or $1{,}185$.

[c] $p < .05$

rate of speaking, *pronunciation*, and *speaking skill*. In both studies, interviewers with lower refusal rates were rated as having higher pitched voices, greater ranges of variation in pitch, greater loudness, faster rates of speaking, and clearer and more distinct pronunciation. *Speaking skill* was only weakly related to refusal rates.

The η^2 for all personal characteristics also differed significantly from zero (Table 3). The largest in Study 1 was for *positive approach* (.21), followed by *general attractiveness*, *age category*, *social class*, and *competence*. The η^2 in Study 2 again were somewhat lower. *General attractiveness*, *competence*, and *age category* had equally high values (.10), followed by *positive approach* and *social class*. In both studies interviewers judged as more generally attractive and, more specifically, as having a more positive approach and greater competence, had lower refusal rates. Interviewers perceived to have higher social class also had lower refusal rates.

Although age was related to interviewer success in both studies, the results were contradictory. It happened that each study included one interviewer noticeably older than the others. In Study 1 she was one of

Table 3. Mean Ratings of Interviewers Grouped by Refusal Rates, Fs, η^2, and Direction of Differences for Personal Characteristic Measures in Two Studies

	Mean Rating: Interviewer Group				
Characteristic	*Low Refusal Rate*	*High Refusal Rate*	F^*	η^2	*Characterization of Interviewers with Low Refusal Rates*
General attractiveness	4.70	3.86	28.06[a]	.15	More attractive
	4.89	*4.28*	*38.15*[b]	*.10*	
Positive approach	5.00	3.93	34.96	.21	More positive
	4.86	*4.29*	*26.52*	*.07*	
Competence	4.67	4.01	16.12	.08	More competent
	5.17	*4.48*	*41.22*	*.10*	
Social class	3.06	2.56	17.82	.10	Higher social class
	3.29	*2.98*	*17.02*	*.05*	
Age category	3.11	2.52	23.84	.11	Younger
	2.65	*3.38*	*110.45*	*.10*	Older

NOTE: Scoring for the *general attractiveness*, *positive approach*, and *competence* measures was reversed from that indicated in Figure 2 for the scales included in them. High scores on these measures indicate positive evaluation. *Social class* was scored so that WORKING = 1 and UPPER = 5. *Age category* was scored so that UNDER 20 = 1 and 60–69 = 6. Italicized figures are for Study 2; other figures are for Study 1.

* For all measures $p < .01$

[a] For characteristics in Study 1 df = 1,85.

[b] For characteristics in Study 2 df = 1,195.

the more successful interviewers, while in Study 2 she was one of the less successful interviewers. We think the results for age reflect vocal and personal characteristics associated with interviewing success, rather than factors related to age per se.

Relation of Vocal Characteristics to Personal Attributions

There was generally a high level of correlation in both studies between mean ratings of vocal and of personal characteristics for the six interviewers (Table 4). Of the vocal characteristics, *variation in pitch* had the highest correlations with interviewer attractiveness, whether measured by *general attractiveness* or two of its components, *positive approach* or *competence*. The other vocal measures also had substantial correlations with the attractiveness measures. Thus, greater attractiveness was associated with voices perceived as relatively high-pitched and loud, and as having a large range of variation in pitch, a fast rate of speaking, clear and distinct pronunciation, and good speaking skills.

Perceptions of higher social class were associated with voices having

Table 4. Correlations Between Mean Ratings of Vocal and Personal Characteristic Measures in Two Studies

	Personal Characteristics				
Vocal Characteristics	*General Attractiveness*	*Positive Approach*	*Competence*	*Social Class*	*Age Category*
Average	.61	.78	.48	.58	.75
pitch	*.56*	*.56*	*.61*	*.63*	*−.93*
Variation	.95	.98	.87	.85	.40
in pitch	*.77*	*.60*	*.94*	*.96*	*−.75*
Loudness	.44	.49	.40	.63	.18
	.67	*.47*	*.91*	*.92*	*−.61*
Rate of	.76	.73	.70	.48	−.15
speaking	*.89*	*.84*	*.71*	*.49*	*−.15*
Pronunciation	.88	.84	.89	.94	.59
	.48	*.41*	*.64*	*.65*	*−.65*
Speaking	.86	.68	.93	.84	−.04
skill	*.53*	*.45*	*.36*	*.19*	*.33*

NOTE. Figures in italics are for Study 2; other figures are for Study 1. All are product moment coefficients based on six interviewers. Measures were scored as described in the notes to Tables 2 and 3.

the same characteristics. Results for age again were contradictory in the two studies, most likely for the reason already noted.

Discussion

All three research goals were met: reliable ratings of vocal and personal characteristics were obtained; vocal characteristics were correlated with desirable and undesirable personal characteristics; and both were related to interviewer refusal rates. In both studies higher pitch, greater variation in pitch, loudness, faster rate of speaking, clearer and more distinct pronunciation, and good speaking skills were all associated with attractiveness (whether determined by the *general attractiveness* measure or by *positive approach* or *competence*) and with higher perceived social class. Interviewers with these vocal and personal characteristics had low refusal rates.[3]

While the associations of vocal and personal characteristics with refusal rates were highly significant statistically, in most instances they were only moderately strong. The interviewers in these studies were a select group, since screening had already eliminated job applicants with voices judged unsuitable, and early departure of some apparently less skilled interviewers excluded them from the studies. A more heterogeneous set of interviewers might yield stronger findings.

---

[3] While a fast rate of speaking for introductory comments may help response rates, in asking questions a slower pace is probably better because it facilitates comprehension and attention to reporting tasks.

Except for the desirability of high average pitch, the findings are consistent with expectations from paralinguistic studies. In these studies (primarily using male speakers) higher pitch was rarely associated with positive attributes. Given gender expectations, higher pitched voices are more acceptable for females just as lower pitched voices are expected of males. It may well be that for females, high pitch (within an acceptable range) is more attractive than low pitch. Although acceptable ranges have not been established, especially high-pitched female voices may not be evaluated positively.

The finding that ratings of loudness were related to success is puzzling. Whereas other acoustical properties of the voices were reproduced accurately in the recordings, vocal intensity was not. The intensities as recorded were roughly the same for all interviewers. The procedure, however, preserved the relative intensities of the various portions of each recording. We can only speculate whether such intensity patterns or some other acoustic properties of the recordings led raters to judge some voices as louder than others.

Findings concerning personal attributions also raise questions. Is there benefit in separating use of a positive approach from competence, or should they be subsumed in a general measure of interviewer attractiveness? Although *positive approach* and *competence* were about equally strongly related to refusal rates in Study 2, *positive approach* was considerably more strongly related than *competence* to refusal rates in Study 1. The concepts are of particular interest since they resemble the task skills and social-emotional dimensions identified in studies of leadership in small groups (Bales, 1966). Further study of them is warranted.

Our studies were limited to female interviewers who were fairly homogeneous with regard to perceived education, social class, age, accent, and use of standard pronunciation. Research is needed to explore the effects of departures from these characteristics as well as to explore whether desirable characteristics differ for male and female interviewers. With regard to the vocal characteristics investigated here, optimal levels have yet to be established. Not only is very high pitch likely to be undesirable, but also a very large range of variation in pitch, great loudness, and an extremely fast rate of speaking.

Finally, we have assumed that vocal characteristics lead to an assessment of the interviewer that in turn disposes the respondent to cooperate or not. However, in our studies vocal characteristics were more strongly related to refusal rates than were personal characteristics. It may be that vocal characteristics have a direct impact on the listener that is not mediated by impressions of the speaker as a person. More research is needed to clarify the process by which these characteristics influence cooperation.

This investigation reinforces the hypothesis that respondents react to cues communicated by the interviewer's voice and may grant or refuse an interview on this basis. Content of the introduction appears to be less important than impressions and affects aroused by the vocalization. To paraphrase a line from *My Fair Lady*, "Respondents don't care what you say, actually; it's whether you say it properly." This should not be surprising to survey researchers, who have always stressed the importance of personal appearance and approach to respondents on the doorstep. Vocal cues, as the only source of information about telephone interviewers, play a particularly significant role in telephone interviewing.

This study should be viewed principally as emphasizing the importance of paralinguistic research to survey methodologists. We hope that other researchers will investigate paralinguistic issues, attempt to replicate our findings, and explore their implications in greater detail. The findings also introduce the question as to whether voice characteristics should be a selection criterion or whether voice training should be added to interviewer training.

References

Addington, D. W.
1968 "The relationship of selected vocal characteristics to personality perception." Speech Monographs 35:492–503.
Allport, G., and H. Cantril
1934 "Judging personality from voice." Journal of Social Psychology 5:37–54.
Apple, W., L. A. Streeter, and R. M. Krauss
1979 "Effects of pitch and speech rate on personal attributions." Journal of Personality and Social Psychology 37:715–26.
Bales, R. F.
1966 "The equilibrium problem in small groups," in A. P. Hare, E. F. Borgatta, and R. F. Bales (eds.), Small Groups: Studies in Social Interactions. New York: Alfred A. Knopf.
Brown, B. L., W. L. Strong, and A. C. Rencher
1973 "Perceptions of personality from speech: effects of manipulations of acoustical parameters." Journal of Acoustical Society of America 54:29–35.
Dillman, D., J. G. Gallegos, and J. H. Frey
1976 "Reducing refusal rates for telephone interviews." Public Opinion Quarterly 40:66–78.
Ellis, D.
1967 "Speech and social status in America." Social Forces 45:431–37.
Giles, H., and P. F. Powesland
1975 Speech Style and Social Evaluation. New York: Academic Press.
Giles, H., W. P. Robinson, and P. M. Smith (eds.)
1980 Language: Social Psychological Perspectives. New York: Pergamon.

Harms, L. S.
1961 "Listener's judgments of status cues in speech." Quarterly Journal of Speech 47:164–68.
Kramer, E.
1963 "Judgment of personal characteristics and emotions from nonverbal properties of speech." Psychological Bulletin 60:408–20.
1964 "Personality stereotypes in voice: a reconsideration of the data." Journal of Abnormal and Social Psychology 68:247–51.
Lalljee, M. G.
1971 "Disfluencies in normal English speech." Unpublished Ph.D. dissertation, Oxford University.
Lay, C. H., and B. F. Burron
1968 "Perception of the personality of the hesitant speaker." Perceptual and Motor Skills 23:951–56.
Mehrabian, A., and M. Williams
1969 "Nonverbal concomitants of perceived and intended persuasiveness." Journal of Personality and Social Psychology 13:37–58.
Miller, N., G. Maruyana, R. Beaber, and K. Valone
1976 "Speed of speech and persuasion." Journal of Personality and Social Psychology 34:615–24.
O'Neil, M. J., R. M. Groves, and C. F. Cannell
1979 "Telephone interview introductions and refusal rates: experiments in increasing respondent cooperation." American Statistical Association meetings, August.
Packwood, W. T.
1974 "Loudness as a variable in persuasion." Journal of Counseling Psychology 21:1–2.
Pear, T. H.
1931 Voice and Personality. London: Chapman and Hall.
Pearce, W. B., and F. Conklin
1971 "Nonverbal vocalic communication and perceptions of a speaker." Speech Monographs 38:235–41.
Putnam, G. N., and E. M. O'Hearn
1955 "The status significance of an isolated urban dialect." Language 31:1–32.
Rosenthal, R.
1973 "Estimating effective reliabilities in studies that employ judges' ratings." Journal of Clinical Psychology 29:342–45.
Sharf, D. J., and M. E. Lehman.
1984 "Relationship between the speech characteristics and effectiveness of telephone interviewers." Journal of Phonetics 12:219–28.
Smith, P.
1979 "Sex markers in speech," in K. R. Scherer & H. Giles (eds.), Social Markers in Speech. England: Cambridge University Press.
Winer, B. J.
1971 Statistical Principles in Experimental Design. 2nd ed. NY: McGraw-Hill.

The Polls:
ABM and Star Wars:
Attitudes Toward Nuclear Defense, 1945–1985

THOMAS W. GRAHAM AND BERNARD M. KRAMER

PRESIDENT REAGAN'S Strategic Defense Initiative (SDI), popularly called Star Wars, is the most recent manifestation of a 40-year debate over developing systems to defend against nuclear attack. This compilation of national survey research is a selection of questions about nuclear defenses taken from our collection of 61 national and 3 state opinion polls conducted from December 1945 through December 1985.[1]

Despite changes in the US-Soviet nuclear balance, several attitudes toward active defenses against nuclear weapons have been remarkably stable. For example, eight out of ten questions which asked about the generic feasibility of developing a defense against nuclear weapons found a majority (up to two-thirds) of the population believing that the U.S. or its scientists could develop such a defense.[2]

Awareness of both the ABM system and the SDI research program has hovered at the 70 percent level both in the 1969 ABM debate and since Reagan's March 1983 speech.[3] However, no more than 16 percent to 24 percent of the public is following the current Star Wars debate "a lot" or "very closely" (1/85, 2/85). In another area concerning awareness, in 1985 only about one-third of the public has heard about the ABM/SALT I treaty (2/85, 9/85).

[1] Neither state polls nor "passive" defenses such as civil defense are covered in this article. For a presentation and analysis of all survey questions and marginals from the 64 surveys, see Thomas W. Graham, Public Attitudes Toward Active Defense: ABM & Star Wars, 1945–1985 (Cambridge, MA: Center for International Studies, MIT, 1986).

[2] These nine polls were conducted in 2/45, 4/46, 8/46, 2/47, 11/49, 6/64, 6/68, 1/85, 5/85. One of the "exceptional" questions asked whether people agreed with expert opinion that it was impossible to defend against surprise attack by rockets carrying atomic bombs (2/45, NORC). The other queried whether the U.S. can "protect" itself from incoming nuclear missiles (5/85, Sindlinger).

[3] Surveys conducted on 4/69, 5/69, 7/69, 4/83, 1/85, 7/85, 9/85.

Thomas W. Graham is a Ph.D. candidate in the Department of Political Science and a Graduate Research Associate at the Defense and Arms Control Studies Program at MIT. Bernard M. Kramer is Professor of Psychology at the University of Massachusetts, Boston.

Public Opinion Quarterly Vol. 50:125–134 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-125/$2.50

Approximately 75 percent of the public holds the (inaccurate) belief that the United States has a fairly effective defense against nuclear weapons (6/64, 6/68, 4/79). In addition, most people are satisfied with our (nonexistent) defense against nuclear attack (5/82, 7/82, 4/84).

In the 1960s, a majority of the population believed that Russia had an existing ABM defense (6/64, 6/68, 4/69). In the 1980s, a plurality of 34 percent thought the Soviets were ahead of the U.S. in Star Wars technology (5/85).

In 1969 deployment of the ABM was favored by the minority of the public who both had heard of the system and had an opinion about building the Safeguard system. In 1971 and 1972, when presented with the idea of spending additional money for the system, support turned to opposition.

A substantial majority of the public has consistently supported the idea of negotiating an ABM agreement or negotiating limitations on Star Wars with the Soviets.[4] Despite this pro-arms control stance, the public believes that moving forward with Star Wars will encourage the Soviet Union to negotiate a nuclear arms control agreement with the U.S. (1/85, 2/85, 5/85).

In the tables that follow, all numbers are percentages. The authors wish to thank the many individuals who helped them discover this wealth of data[5] and the survey organizations that graciously have made it available for academic research:

ABC/WP—ABC & Washington Post
Abt—Clark Abt Associates
AIPO—American Institute of Public Opinion/Gallup
CBS/NYT—CBS & New York Times
Harris—Louis Harris & Associates
LAT—Los Angeles Times
MIS/N—Market Information Service & Dr. Jiri Nehnevaja
M&K—Marttila & Kiley
NORC/N—National Opinion Research Center & Dr. Jiri Nehnevaja
ORC/N—Opinion Research Company & Dr. Jiri Nehnevaja
P&S—Penn & Schoen Associates
Sind—Sindlinger & Associates

ABM

NORC/N & MIS/N [respondents ranked the desirability of building an ABM from +3 to −3]:

	−3	*−2*	*−1*	*0*	*+1*	*+2*	*−3*	*DK*
12/63	8	3	2	5	8	16	59	—
6/64	5	2	1	4	6	13	67	1
2/66	7	1	2	4	7	9	69	1
3/72	6	4	4	14	14	16	43	5

[4] Surveys were conducted on 6/68, 8/70, 6/71, 2/72, 3/72, 6/72, 6/73, 11/73, 4/84, 6/84, 7/84, 11/84, 1/85.

[5] Special thanks to the always cooperative Marilyn Potter, Lois Timms and John Benson of the Roper Center for Public Opinion Research at the University of Connecticut; Ann Gray and Josephine Marsh of the Louis Harris Data Center, Institute for Research in Social Sciences at the University of North Carolina; Professor Jiri Nehnevaja and Nicholas R. Trio of the University Center for Social and Urban Research at the University of Pittsburgh; Tom W. Smith and Pat Bova at NORC; and Professor Jack P. Ruina of MIT for providing moral support and financial aid from a Carnegie Corporation grant.

NORC/N [respondents ranked likelihood that an ABM would be built from 0 to 10]:

	0	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*DK*
6/64	4	3	3	3	4	15	11	10	12	11	23	1
2/66	11	7	6	6	7	20	9	7	8	8	12	1

NORC/N [respondents ranked desirability of deploying an ABM from +3 to −3; response options abbreviated here]:

6/1964	*+3*	*+2*	*+1*	*0*	*−1*	*−2*	*−3*	*DK*
Around all large cities	61	17	9	2	2	1	6	1
Around your city	46	16	14	7	3	3	6	1
If risk fired by accident	29	15	15	4	7	6	9	1
If shoot down 9 of 10	67	14	8	2	1	1	4	1
If shoot down 1 of 3	42	19	17	4	4	3	4	1

NORC/N: Now I'll read a few more statements to you. Please tell me whether you strongly agree, agree, disagree, or strongly disagree with each statement, or whether you are undecided.

6/1964	*Strongly Agree*	*Agree*	*Disagree*	*Strongly Disagree*	*Undec*
If we have such missiles around our cities, we should have shelters to protect against fallout because some enemy weapons will get through the defense anyway.	18	65	7	2	8
Such missiles will make the Russians think that we are going to start a war; therefore, they might start one.	3	21	53	8	13
Antimissile missiles will lead to a stepping up of the arms race.	4	44	33	2	16
Such missiles will cost too much money to be worth while.	3	14	58	10	17
New offensive weapons will soon put antimissile missiles out of date, so it really is not worth having them.	1	12	53	7	25
Putting in antimissile missiles makes it more difficult to reach agreements on arms control and disarmament.	2	28	47	4	18
Antimissile missiles will make America stronger, and an enemy will be even less likely to attack us than without these missiles.	15	61	13	2	8

NORC/N [respondents evaluated types of cities that should be defended; q abbrev.]:

6/1964	*Most*	*Next Most*	*Least*	*Next Least*
Cities with the largest population	19	12	4	10
Cities with the most industry	24	36	3	4
Cities with military establishments around them	43	25	3	3
Cities that are important in American history	1	2	48	32
Sea and land transportation centers	9	20	7	9
Cities which are drawn by lot to insure that all have equal chances to be defended	2	2	33	39

ORC/N [respondents evaluated desirability of ABMs under certain circumstances; q abbrev.]:

6/1968	*−3*	*−2*	*−1*	*0*	*+1*	*+2*	*+3*	*DK*
How desirable is it to put ABMs around all large cities in America?	5	3	3	9	9	11	48	11
How desirable is it if they only could shoot down about one out of every two missiles?	22	7	5	9	9	10	23	16
A bigger program costing about 40 billion dollars has been suggested. How desirable is such a program to you?	14	4	5	16	9	8	26	17
If we know that the Russians have an active ABM program, how desirable would it be that the U.S. have one?	3	1	1	5	5	9	66	9

ORC/N: Here are some possible consequences of this ABM program. I would like to know for each one whether you strongly agree, agree, disagree, or strongly disagree that it might be a consequence.

6/1968	*Strongly Agree*	*Agree*	*Disagree*	*Strongly Disagree*	*DK*
Having ABMs around our cities will make people think that war is more likely	13	37	36	6	8
If we have ABMs, the Russians will think that we are going to start a war	5	25	50	9	11
Having ABMs will make America stronger and an enemy will be less likely to attack	26	48	15	3	9
New offensive weapons will put ABMs out of date, so it is really not worth having them	4	17	43	9	27

HARRIS [respondents asked to evaluate various defense options; q abbrev.]:

12/1968	*Ought to Be Done*	*Should Not Do*	*Not Sure*
Build up a system of antimissile defenses.	60	23	17
Convert the space program into a system of nuclear weapons space stations	25	49	26

AIPO [asked of those who had heard of the ABM]: Do you happen to have an opinion about the ABM program as submitted to Congress by President Nixon?

	Yes	*No*	*Not Heard*
4/69	40	27	33
5/69	41	31	28
7/69	42	27	31

AIPO [asked of those who had heard of the ABM and had an opinion]: Do you favor or oppose the ABM program submitted by Nixon?

	Favor	*Oppose*	*Not Heard or No Opinion*
4/69	25	15	60
5/69	25	14	61
7/69	23	18	59

HARRIS [prefatory description deleted]. Do you tend to approve or disapprove of President Nixon's decision to go ahead with the antimissile system?

4/1969 *Approve* (48%) *Disapprove* (25%) *Not Sure* (27%)

HARRIS [Respondents asked to agree/disagree with statements]:

4/1969	*Agree*	*Disagree*	*Not Sure*
Many scientists think the system is a mistake because it could not prevent the loss of millions of U.S. lives in an atomic attack.	40	28	32
By taking this step, President Nixon hurt the chances of reaching arms control agreements with the Russians.	19	52	29
We could have used the seven billion dollars better for our education, health, housing, and poverty needs at home.	49	31	20
7/1969			
The Safeguard ABM anti-missile system is necessary in order for us to keep up with the Russians in the nuclear arms race.	50	24	26
4/1969 and 7/1969			
It is better to be overprepared militarily than to be caught short without proper defenses.			
4/1969	78	9	13
7/1969	84	7	9
4/1969 and 7/1969			
In a nuclear age, the U.S. should concentrate on peaceful arms control rather than building more missile systems like the Safeguard ABM.			
4/1969	47	29	24
7/1969	50	24	26

Star Wars

SIND: If the U.S. had the capability of changing this situation [that we have no defense against ballistic missiles] by deploying an antiballistic missile defense, would you favor it being done?

	Yes	*No*	*No Opinion*
8/1982	86	10	4

HARRIS [respondents asked to agree/disagree with statements]:

While it might be possible to develop such defensive systems in outer space, once such means of destruction were built, they could easily be turned into threatening mankind with new and frightening space wars.

	Agree	*Disagree*	*Not Sure*
4/1983	72	24	4

Once the Russians knew we were building a new outer space antinuclear defense system, they would be more willing to negotiate a treaty that would halt the nuclear arms race.

4/1983	40	55	5

The only way to avoid a nuclear war is to develop new weapons in space that can shoot down all nuclear missiles.

	38	57	5

All in all, do you favor or oppose spending billions of dollars for the U.S. to develop a laser-beam and particle-beam outer space defense system?

4/1983	36	58	5

ABC/WP (asked of 65% who had heard/read of Reagan SDI plan): Do you favor or oppose developing such defensive weapons, or what?

	Favor	*Oppose*	*No Opinion*
4/1983	54	37	8

ABC/WP: Just your best guess, if the United States does develop such defensive weapons, would that increase or reduce the arms race between the United States and the Soviet Union, or what?

	Increase	*Decrease*	*Not Change* (vol)	*No Opinion*
4/1983	57	24	9	10

CBS/NYT: President Reagan has proposed developing a defensive system that would destroy incoming Russian missiles before they reach the United States. Some people say it might be difficult technologically, but we should try to develop it. Other people say it would be impractical, expensive and sounds like science fiction. Do you think we should try to develop the system, or not?

	Should	*Should Not*	*Don't Know*
4/1983	67	25	8

P&S: Do you favor or oppose plans to develop defensive weapons that would operate in space in order to protect the U.S. by destroying any incoming missiles?

	Favor	*Oppose*	*Don't Know*
4/1984	75	17	8

HARRIS [preliminary wording about U.S.–Soviet negotiating differences deleted here]:

Do you favor or oppose the U.S. and the Soviet Union sitting down to negotiate an agreement to outlaw the use of weapons in outer space?

	Favor	*Oppose*	*Don't Know*
7/1984	82	16	2

HARRIS: How optimistic are you that the U.S. and Russia will come to an agreement anytime soon to outlaw the use of weapons in space—very optimistic, somewhat optimistic, not very optimistic, or not optimistic at all?

	Very Optimistic	*Somewhat Optimistic*	*Not Very Optimistic*	*Not Optimistic at All*
7/1984	9	23	36	31

ABT: Would you support a U.S. attempt to build a defensive system against nuclear missiles and bombers under the following conditions:

If it did not use nuclear explosives and was only partially effective such as being able to shoot down 90% of all Soviet nuclear missiles and bombers, and letting 10% get through?

9/1984 *Yes* (57%) *No* (35%) *Don't Know* (8%)

If it used nuclear explosives and was only partially effective such as being able to shoot down 90% of all Soviet nuclear missiles and bombers, and letting 10% get through?

9/1984 *Yes* (44%) *No* (48%) *Don't Know* (8%)

If it was expensive and increased defense spending by at least 10% and increased government spending by 3%?

9/1984 *Yes* (58%) *No* (35%) *Don't Know* (7%)

CBS/NYT: Some people have proposed a defense system in space to protect the United States from nuclear attack, the concept called "Star Wars." Do you think this would make us more secure, or would it just speed up the arms race?

	Make More Secure	*Speed Up Arms Race*	*Both* (vol)	*Neither* (vol)	*No Opinion*
10/1984	31	48	3	2	16

CBS/NYT: Ronald Reagan has proposed developing a defensive nuclear system in space that would destroy incoming missiles before they reach the United States, a system some people call "Star Wars." Do you think such a system could work?

1/1985 *Yes* (62%) *No* (23%) *No Opinion* (15%)

CBS/NYT: Do you think that developing this system would make negotiations with the Soviet Union easier?

1/1985 *Yes* (48%) *No* (41%) *No Opinion* (11%)

CBS/NYT: Do you think this would be worth the amount of money it would cost?

1/1985 *Yes* (40%) *No* (46%) *No Opinion* (14%)

CBS/NYT: Would having this system make you feel secure or would putting nuclear weapons in space worry you?

1/1985 *Make Feel Secure* (25%) *Worry You* (60%) *Both* (vol. 4%)
Neither (vol. 2%) *No Opinion* (8%)

ABC/WP (asked of 63% who had heard/read of Reagan SDI proposal): Do you favor or oppose developing such defensive weapons, or what?

1/1985 *Favor* (49%) *Oppose* (44%) *No Opinion* (7%)

LAT: Generally speaking, do you favor or oppose an agreement to outlaw the use of all military weapons in outer space? Is that (favor/oppose) strongly or somewhat?

1/1985 *Favor Strongly* (43%) *Favor Somewhat* (16%)
Oppose Somewhat (15%) *Oppose Strongly* (16%)
Not Sure (10%)

AIPO: How closely have you followed the discussions over the administration's so called Star Wars proposal—that is, its proposal to develop a space-based defense against nuclear attack?

2/1985 *Very Closely* (16%) *Fairly Closely* (51%) *Not Closely at All* (30%)

AIPO: Would you like to see the United States go ahead with the development of such a system or not?

2/1985 *Yes* (52%) *No* (38%) *No Opinion* (10%)

AIPO: In your opinion, would the United States' developing this system increase or decrease the likelihood of reaching a nuclear arms agreement with the Soviet Union?

2/1985 *Increase* (47%) *Decrease* (32%) *No Difference* (vol. 13%) *No Opinion* (8%)

HARRIS [preliminary description of Reagan program deleted here]: Once the Russians knew we were building a new outer space antinuclear defense system, they would be more willing to agree to a treaty that would halt the nuclear arms race.

3/1985 *Agree* (52%) *Disagree* (44%) *Not Sure* (4%)

HARRIS: The only way to avoid a nuclear war is to develop new weapons in space that can shoot down all nuclear missiles.

3/1985 *Agree* (44%) *Disagree* (52%) *Not Sure* (4%)

HARRIS: Once it looked as though we were capable of defending against today's nuclear weapons, the Soviets would then go all-out to develop new kinds of nuclear and other weapons we couldn't defend against.

3/1985 *Agree* (75%) *Disagree* (20%) *Not Sure* (5%)

HARRIS: Even though President Reagan wants to spend only $3.8 billion at first on developing a laser defense system, the cost could rise to $25 billion to $50 billion on a system that might not even work.

3/1985 *Agree* (63%) *Disagree* (30%) *Not Sure* (7%)

HARRIS: All in all, do you favor or oppose spending billions of dollars for the U.S. to develop a laser-beam and particle-beam outer space defense system?

3/1985 *Favor* (39%) *Oppose* (56%) *Not Sure* (5%)

SIND: Some people say that in the development of any strategic defense system that could destroy incoming missiles, the Soviet Union is far ahead of the United States, while other people are saying that the United States is far ahead of the Soviet Union. What do you think?

	Soviet Union Ahead	*Soviet Union Behind*	*Both the Same*	*Not Sure*
5/1985	34	26	17	23

ABC/WP (asked of those 41% who approved SDI): Currently the U.S. and the Soviet Union have an antiballistic missile treaty that prohibits both nations from developing certain weapons. Suppose the U.S. had to violate or abandon that treaty in order to develop the space-based weapons. Would you still favor the development of those space-based weapons, or not?

	Yes	*No*	*Don't Know*
7/1985	63	32	5

M&K: The Reagan administration is now working on this program known as Star Wars. It will attempt to build a new defensive system in outer space that could shoot down nuclear missiles fired at the U.S. Since the program is currently only a research project, it is impossible to predict how complete a defense it will provide. I'm going to read a list of four possible Star Wars systems. For each one, tell me if you would strongly support, support, oppose, or strongly oppose building each type of system.

9/1985	*SS*	*S*	*O*	*SO*	*NS*
A system that was perfect and could successfully defend against all incoming nuclear weapons.	58	28	5	4	4
A system that could protect our missile sites and some population centers, but could not guarantee the safety of many of our major cities.	12	32	34	16	6
A system designed only to protect U.S. missiles, key military bases, and Washington, D.C., but not other areas.	5	16	39	34	5
A system that could provide a complete defense against long-range nuclear missiles, but cannot defend against missiles fired from submarines or bombers.	8	32	34	17	10

M&K: Now I'm going to read several statements people have made about the Star Wars proposal. Some of the statements argue in favor of Star Wars and some argue against it. Regardless of how you feel, please listen to each statement and tell me whether you strongly agree, mostly agree, mostly disagree, or strongly disagree with each one.

9/1985	*SA*	*MA*	*MD*	*SD*	*NS*
If we build the Star Wars system the Soviets will just start building more and more weapons until they can penetrate it.	37	39	13	5	6
Star Wars is the kind of bold, new step we need to lead us away from the arms race.	18	29	26	17	11
Star Wars will make nuclear weapons impotent and obsolete.	11	22	32	22	14
It would be a dangerous mistake to expand the nuclear arms race into outer space.	27	27	25	12	4
Star Wars will encourage the Soviet Union to reach an agreement with us to reduce nuclear weapons.	16	31	27	18	9
The Soviet leaders are extremely concerned about Star Wars, which means we must be on the right track with this idea.	28	36	18	8	10
Even if it's a good idea, Star Wars costs too much to seriously consider building, especially at a time when our federal deficit is already at an all-time high.	29	29	21	11	9

P&S: If a nuclear freeze were negotiated, should development of such defensive space weapons continue or should development be abandoned?

4/1984 *Continue* (54%) *Abandon* (35%) *Don't Know* (11%)

LAT: Some people say that the first thing we have to do is conduct Star Wars research to find out if the idea works. Then when we know what is at stake, there will be time enough to negotiate with the other side. But other people say that, just as soon as one side starts research, the other side will try to improve its weapons so they can avoid Star Wars defenses, and nuclear development escalates. Do you think the two sides should wait until research turns up something before they negotiate or do you think the research should be banned at the start to avoid nuclear weapons escalation?

1/1985 *Ban* (55%) *Delay* (33%) *Not Sure* (11%) *Refused* (1%)

LAT: The idea behind our nuclear defense over the past 40 years is called "Mutual Assured Destruction." It says that what keeps both the United States and the Soviet Union from firing off their nuclear weapons at each other is the certainty that the other side would counterattack with dreadful destruction. Now some people think the "Star Wars" system would destabilize this balance of terror. They claim that if one side had a way to defend itself against nuclear counterattack, it could launch an aggressive war without fear of retaliation. Do you think that a "Star Wars" system threatens its opponents or not?

1/1985 *Yes* (56%) *No* (28%) *Not Sure* (15%) *Refused* (1%)

AIPO: In your opinion, would developing this system make the world safer from nuclear destruction, or less safe?

2/1985 *Safer* (50%) *Less Safe* (32%) *No Difference* (vol 11%) *No Opinion* (7%)

ABC/WP: Have you read or heard about plans by the Reagan administration to develop weapons in outer space that could destroy nuclear missiles fired at the United States by the Soviet Union or other countries? (84% responded yes) Reagan calls the research on these weapons SDI, for Strategic Defense initiative, and some people refer to it as "Star Wars."

Supporters say such weapons could guarantee protection of the United States from nuclear attack and are worth whatever they cost. Opponents say such weapons will not work, will increase the arms race, and that the research will cost many billions of dollars. How about you: would you say you approve or disapprove of plans to develop such space-based weapons?

7/1985 *Approve* (41%) *Disapprove* (53%) *Don't Know* (5%)

M&K: Supporters of the ABM treaty say it has made the world safer by giving both the United States and the Soviet Union the security of knowing that the other side cannot use nuclear weapons without being destroyed in retaliation. At some point, development of Star Wars will require the U.S. to either withdraw from the ABM treaty or violate the treaty. If you had to choose between developing Star Wars or keeping the ABM treaty, which would you choose?

9/1985 *Develop Star Wars* (37%) *Keep ABM Treaty* (48%) *Not Sure* (15%)